Title
Essays in Development Economics

Author
Bazzi, Samuel Ali

Publication Date
2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA, SAN DIEGO

Essays in Development Economics

A dissertation submitted in partial satisfaction of the
requirements for the degree
Doctor of Philosophy

in

Economics

by

Samuel Ali Bazzi

Committee in charge:

Professor Gordon Hanson, Chair
Professor David Fitzgerald
Professor Craig McIntosh
Professor Paul Niehaus
Professor James Rauch

2013

The dissertation of Samuel Ali Bazzi is approved, and it is acceptable in quality and form for publication on microfilm and electronically:

Chair

University of California, San Diego

2013

DEDICATION

To my beloved parents, Carolyn and Ali, my brother Joseph and dearest Angela.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Professor Gordon Hanson whose constant support and encouragement has been invaluable.

I am also extremely grateful for the help and advice from my other committee members, Craig McIntosh, Paul Niehaus, and James Rauch. David FitzGerald provided helpful feedback on Chapter 1. I also benefited from numerous conversations with Marc Muendler and Prashant Bharadwaj.

I would also like to acknowledge my coauthors on the chapters in this dissertation—Asep Suryahadi, Sudarno Sumarto, and Michael Clemens—as well as those in other ongoing and completed projects including Steve Radelet, Rikhil Bhavnani, Arya Gaduh, Alex Rothenberg, and Maisy Wong.

I would also like to thank my peers, Stephen Morris, Jacob Johnson, Jong Myun Moon, and Jedrzej Zieleniak, for many helpful conversations over the years.

Chapter 1, in part is currently being prepared for submission for publication of the material, Bazzi, Samuel. The dissertation author was the primary investigator and author of this material.

Chapter 2, in part, has been submitted for publication of the material, Bazzi, Samuel; Suryahadi, Asep; Sumarto, Sudarno. The dissertation author was the primary investigator and author of this paper.

Chapter 3, in part, has been published as it appears in American Economic Journal: Macroeconomics, 2012. Bazzi, Samuel. The dissertation author was the primary investigator and author of this paper.

VITA

2005	B.A. in Economics and International Relations, University of Southern California
2008-2013	Teaching Assistant, University of California, San Diego
2013	Ph.D. in Economics, University of California, San Diego

PUBLICATIONS

Bazzi, Samuel and Michael A. Clemens, "Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth", *American Economic Journal: Macroeconomics*, 2013, 5(2): 152–186.

Clemens, Michael A., Steven Radelet, Rikhil Bhavnani, "Counting Chickens When They Hatch: Timing and the Effects of Aid on Growth," *The Economic Journal*, 2012, 122(561): 590–617.

ABSTRACT OF THE DISSERTATION

Essays in Development Economics

by

Samuel Ali Bazzi

Doctor of Philosophy in Economics

University of California, San Diego, 2013

Professor Gordon Hanson, Chair

This dissertation explores three topics in economic development. Chapter 1 develops and tests a framework for identifying the effect of changes in individual incomes on aggregate migration. Using rich data from Indonesia and a novel two-step estimator suggested by the theory, I provide evidence that financial constraints limit international labor mobility from low-income settings. Chapter 2 uses a large-scale cash transfer program in Indonesia to demonstrate the importance of timing and expectations in evaluating fiscal interventions. Overall, the study finds consumption smoothing behavior similar to that found in evaluations of cash transfer programs in rich countries. Chapter 3 offers a constructive methodological critique of the econometric methods used to identify the causal determinants of economic growth.

Chapter 1

Wealth Heterogeneity, Income Shocks, and International Migration: Theory and Evidence from Indonesia

Abstract

This paper investigates the extent to which financial constraints limit international labor migration flows in a developing country context. Income growth in these settings can have two countervailing effects. Rising income may relax liquidity constraints that prevent profitable migration among poor households. However, higher income also implies smaller wage gaps with rich countries and a higher opportunity cost of migrating. Although acknowledged in the literature, existing theories make it difficult to disentangle these offsetting effects empirically. I solve this difficulty by incorporating wealth heterogeneity, fixed migration costs, and different types of income shocks into a microfounded model of village-level migration flows. Analytic solutions for village migration rates are obtained by exploiting observable heterogeneity in household landholdings, which follow a Pareto distribution. I then test for financial constraints in Indonesia using new administrative panel data on international migration from 66,000 villages. I capture income variation arising from transitory rainfall shocks and a large, sustained increase in domestic rice prices following an unanticipated ban on rice imports. Using a two-step estimator suggested by the theory, I find that positive agricultural income shocks are associated with significant increases in the share of village residents working abroad, particularly in villages with a greater mass of small landholders. Migration flows are more responsive to rainfall shocks in underdeveloped villages—in terms of bank presence and other measures of baseline wealth levels. The empirical findings are consistent with the prevalence of binding financial constraints to labor mobility. I rule out aggregation bias using auxiliary survey data. Using the structural model, I estimate village-specific migration costs and show how financial constraints imply large interregional differences in potential future emigration flows.

1.1 Introduction

Every year, several million individuals from developing countries migrate abroad for work. Millions of others in the labor force aspire to do the same.[1] These migrants typically realize substantial income gains for themselves and their families remaining at home (see Clemens, 2011). Moreover, recent studies identify sizable global welfare and efficiency gains to greater international migration, particularly from low-income countries (Benhabib and Jovanovic, 2012; Kennan, forthcoming; Klein and Ventura, 2009). Yet, barriers to labor mobility are often pervasive in these settings. Unlocking the large potential economic benefits of international migration highlighted in existing literature requires a deeper understanding of the origin, magnitude, and prevalence of such barriers.

Even in the absence of policy and informational barriers to mobility, the

[1] A recent Gallup/IOM (2011) world poll across 150 countries estimates that around 1.1 billion individuals would like to work outside their home countries at least temporarily. The vast majority reside in developing countries.

costs of migration may exceed the financial means of relatively poor households. The inverted U relationship between emigration rates and national income observed in cross-country historical data from Europe has been posited as evidence of this type of liquidity constraint.[2] However, there remains disagreement over what this pattern implies today about the effects of rising income in developing countries on international migration flows (see Hatton and Williamson, 2011). Standard theories in which migration costs are implicitly financed through past savings or by borrowing against future earnings cannot directly inform this debate; in these models, expected wage differentials are sufficient to identify migration outcomes (e.g., Borjas, 1987). Recent microeconomic models relax this assumption and find new implications for the self-selection of migrants in a cross-sectional context (McKenzie and Rapoport, 2007; Orrenius and Zavodny, 2005). However, additional inputs and structure are needed to test for the importance of financial constraints in shaping migration flows.

In an ideal research setting, one would randomly relax income constraints in the population of potential international migrants and examine subsequent emigration flows across the baseline wealth distribution. Such an experiment has not yet been conducted. To move ahead, we therefore need to know whether other exogenous shocks can be used to test for the prevalence of financial barriers. But this is not straightforward since positive income shocks at home not only relax liquidity constraints but also affect relative prices of labor and may increase the opportunity cost of migration, which makes identification difficult.[3]

In this paper, I develop and test a theoretical framework that clarifies how to disentangle these offsetting effects of income shocks. The microfounded model characterizes aggregate migration flows from rural villages, where (i) landholdings are a key source of income-generation and wealth, and (ii) important covariate income shocks are observable in the form of rainfall and agricultural commodity prices. In this setting, the (land-)poor may be unable to afford to migrate and the

[2]That is, emigration rates are lowest among the poorest and wealthiest countries (see Hatton and Williamson, 1998).

[3]Throughout this paper, in referring to (agricultural) income, rainfall, or price *shocks*, I have in mind *positive shocks*.

(land-)rich may lack the incentives to do so. This could give rise to the inverted U relationship between international migration and wealth posited in prior studies. Here, I show how this cross-sectional relationship affects the response of migration flows to income shocks at home. In the presence of cash-in-advance constraints, positive productivity or producer price shocks may enable lower landholding individuals to pay the fixed costs of undertaking profitable migration opportunities. Relatively higher landholding individuals for whom migration was previously both profitable and affordable may subsequently remain at home to take advantage of higher returns to agriculture. Which effect dominates—and hence the change in aggregate migration rates—depends on the distribution of landholdings and the prevalence of liquidity constraints among potential migrants.

My theoretical approach has several advantages relative to existing literature. First, I borrows insights from heterogeneous firm trade theory (Melitz, 2003) to derive analytic and empirically-founded expressions for village migration rates. Helpman et al. (2008) aggregate firm-level exports into international trade flows using unobservable firm heterogeneity in productivity. Here, I show how observable household heterogeneity in productivity can be used to aggregate individual migration choices into migration rates. In particular, I document and then exploit the stylized fact that landholdings follow a Pareto distribution, allowing the dispersion parameter, λ, to vary across villages.[4] The income thresholds within which migration is both affordable and profitable may then also differ across villages depending on fixed migration costs and economic conditions at home, without sacrificing tractability.

An appealing feature of this approach is that it rationalizes zero migration rates as a possible equilibrium outcome. Zeros are a common feature of migration datasets that cannot be explained using standard random utility models.[5] The

[4]Although other studies examine the relationship between landholdings and migration (Halliday, 2006; Jayachandran, 2006; Mendola, 2008; Meng, 2009, 2010; VanWey, 2005), the present study is, to my knowledge, the first to utilize the distribution of landholdings as a means of aggregating individual migration choices over subgroups—in this case, villages—within the population.

[5]In the widely used Global Migrant Origin Database, for example, over 25 percent of 226×225 bilateral pairs of countries have no foreign-born individuals based on the 2000 round of Population Censuses; Beine et al. (2011) report zeros in approximately 35 percent of bilateral pairs in their

theory developed here distinguishes between an extensive margin of any village residents working abroad and an intensive margin migration rate conditional on having any migrants. Using the Pareto formulation, the location of the wealth distribution affects the extensive margin, and the shape of the distribution affects the intensive margin. This distinction allows barriers to mobility in villages with a long history of migration to differ from those in villages with no recent migration history. Thus, income and other shocks may have different effects on the two margins, which is consistent with the theory of endogenous migration costs based on network externalities (Carrington et al., 1996).

Finally, I offer a convenient way to identify liquidity constraints that is consistent with the permanent income hypothesis and does not require modeling endogenous supply-side financial institutions or social networks. I do so by distinguishing between transitory rainfall shocks and potentially permanent commodity price shocks. In the absence of liquidity constraints among potential migrants, rainfall shocks should have no effect on migration decisions whereas positive price shocks should reduce migration flows, and larger reductions should occur in villages with greater mass of large landholders. However, when liquidity constraints are binding, the model predicts a positive effect of both types of shocks on village migration rates, and these effects are larger in villages with a greater mass of small landholders. Intuitively, positive shocks increase own household liquidity but may also loosen thin informal credit markets.

Guided by the theory, I test for the prevalence of financial constraints in Indonesia using new administrative panel data on international labor migration from nearly 66,000 villages. I capture income variation arising from rainfall shocks and a large, exogenous increase in the domestic price of rice, Indonesia's most important agricultural product.[6] In January 2004, in the midst of pre-election campaigning, the government unexpectedly banned rice imports, which, although small rela-

study of diasporas using data from 1990 and 2000; roughly 95 percent of origin country $\times$ destination US state observations are zero in Simpson and Sparber (2010).

[6]Similar strategies have been pursued elsewhere in the development economics literature. Edmonds and Pavcnik (2003) exploit abrupt rice price shocks after Vietnam's trade liberalization to study the effect of income shocks on child labor, and Qian (2008) uses agricultural policy reforms in China to study the effect of differential changes in the returns to male and female labor on gender ratios.

tive to national production, stabilized domestic prices historically. Indonesian rice producers effectively operated in autarky over the next several years, and in late-2005, domestic prices began a steep ascent, eventually surpassing historical peaks (Dawe, 2008). By late-2007, the real domestic price was over 50 percent higher than in January 2004 and more than 30 percent above the world price (Figure 1.1). Moreover, the magnitude of the shock varied considerably across regions, and I am able to use this additional variation for reasons discussed in Section 1.4. Lastly, I capture predetermined heterogeneity in household landholdings using data from a universal agricultural census in 2003.

The empirical results suggest that financial constraints are an important barrier to international migration. First, in reduced form specifications, positive rainfall and rice price shocks are associated with an increase in international migration rates between 2005 and 2008. However, because these estimates conflate the extensive and intensive margins, they are uninformative about the theoretical mechanisms and, in particular, the extent to which liquidity constraints bind on either margin. Taking an alternative two-step approach suggested by the theory, I find on the intensive margin that rainfall and rice price shocks lead to significant increases in the share of village residents working abroad. Furthermore, the elasticity of flow migration rates with respect to these shocks is higher in villages with a greater mass of small landholders. On the extensive margin, villages with larger maximum landholdings are more likely to have any migrants. Both findings are consistent with binding financial constraints in the model.

Other evidence supports a financial constraints interpretation. First, positive rainfall shocks have smaller effects in more economically developed and relatively wealthier villages, in terms of deeper bank presence, higher initial mean household expenditure per capita, more widespread access to technical irrigation, and higher aggregate productivity of land. Second, the positive effects of rice price shocks are largest and most precisely estimated for landholdings distributions specific to rice production. If migration choices were financially *un*constrained, then we should observe more muted effects among rice producers (captured by these distributions), for whom income at home and the opportunity cost of migration are rising faster. These heterogeneous effects of the rice price shock are also consistent with non-rice agricultural households experiencing a negative real income shock rendering migration infeasible among small, financially-constrained landholders. Moreover, all empirical results are conditional on the share of (near-)landless net rice consumers.

I also demonstrate a tight link between the underlying, heterogeneous household migration choices and the estimated village-level elasticities of migration flows with respect to income shocks. Using auxiliary household panel data, I first show that positive rainfall and price shocks increase the probability of having any international labor migrants, especially for households with small landholdings. I then aggregate the household migration elasticities over the village-specific Pareto landholdings distributions and obtain implied village-level elasticities that are very close to the actual elasticities estimating using the main village-level data. This finding provides evidence against aggregation bias and in support of the theory. However, the village-level regression elasticities exhibit much more variation because, in line with the theory, household migration elasticities at a given landholding size are allowed to vary across villages depending on fixed migration costs. Such variation cannot be fully captured in household survey data that is not representative at the village level.

Finally, I recover lower bounds on village-specific migration costs using a simple calibration exercise. The two-step econometric model makes it possible to estimate potential migration rates and hence recover costs for all villages including those with no migrants over the sample period. The implied costs for prevailing two-year work contracts in destination countries range from 400-5100 USD and equal 1.5 years of total household expenditures in the median rural household. Costs vary systematically across the country and are lower in villages (i) located closer to peninsular Malaysia, Singapore, or emigration hubs within Indonesia, or (ii) with longstanding networks in Malaysia. The implied net income for an individual working abroad is nearly 40 percent larger than cumulative expenditures in the median rural household over a typical two year period. This differential exceeds 100 percent in many areas of densely populated rural Java and remote regions of Eastern Indonesia.

It is important to note that the main empirical results apply to the *average* Indonesian village rather than the typical potential migrant per se. In a heterogeneous developing country (or village) population, migration may be both an investment subject to liquidity constraints and a means of overcoming liquidity constraints to other investments. The findings here shed new light on the prevalence of the former but are not inconsistent with the latter, which is addressed extensively in existing literature (see Yang, 2011).

This paper is related to a large literature looking at the relationship between income growth and migration. Although some micro-level studies allow the effect of income shocks on migration choice to vary with wealth, these reduced form relationships have not been convincingly connected via theory to a test for liquidity constraints. For example, Jayachandran (2006) finds that transitory rainfall shocks increase rural to urban migration in India but only among small landholders.[7] Here, I show how to aggregate these household responses, differentiate between transitory and persistent income shocks, and characterize the importance of liquidity constraints for explaining overall migration flows. Meanwhile, several macro-level studies regress bilateral migration flows on measures of average income and inequality in sending and receiving countries. Such studies commonly find a null coefficient on sending country income levels or shocks (see Gaston and Nelson, 2013, for a review). Here, I develop a tractable framework using the wealth distribution at home to identify both the sign and magnitude of the change in migration rates subsequent to different types of income shocks at home. This framework can explain how the liquidity and opportunity cost effects of income shocks can offset each other in macro data, thus solving the difficulty of aggregating individual responses to shocks that affect both the relative returns to and ability to finance a costly investment in migration.

By using wealth heterogeneity in this way, I am also able to go beyond identifying liquidity constraints among the poorest households allowing one to

[7] For example, earthquakes have different effects on Salvadoran households depending on landholdings in Halliday (2006) or credit access in Yang (2008b). In Ardington et al. (2009), cash transfers have differential effects on South African households depending on socioeconomic status.

characterize the effects of financial barriers on aggregate migration flows. I do so by extending landed vs. landless dichotomy in Jayachandran's (2006) model and by allowing for origin-specific fixed costs. A few recent studies identify liquidity-constrained migration choice using randomized cash transfers to the poorest households in Mexico (Angelucci, 2012) and Bangladesh (Bryan et al., 2012).[8] Their approaches are ideal for identifying binding liquidity constraints among individuals with very low levels of wealth. However, without stratification along the entire wealth distribution and transfers that vary (monotonically) with wealth and income-generating opportunities at home, such an approach cannot

More broadly, I advance a new microfounded approach to estimating models of international migration flows. In particular, I clarify how to overcome the aggregation bias that arises when estimating the elasticity of migration flows with respect to income shocks in the presence of localized origin-specific fixed migration costs.[9] By comparing international emigration rates across villages within a single developing country, I am able to hold destination country policy barriers constant while testing for the importance of other barriers to labor mobility that vary across regions within the origin country but have implications for aggregate migration. This approach opens the door to future research aimed at reconciling the gap between migration flows observed in aggregate data and those predicted on the basis of international wage differentials alone. In an analogous manner, recent studies in the trade literature use sector-level data to demonstrate how imperfect credit markets can reduce aggregate international trade flows (e.g., Manova, forthcoming).

Finally, this paper is relevant to three additional areas of the literature. First, my theory not only makes novel use of analytical advances in recent trade literature but also relates to a broader literature on heterogeneity and aggrega-

[8]Other prominent shocks used in the non-experimental literature include, among others, inheritance laws (Abramitzky et al., 2012), deep social networks (Borger, 2010; McKenzie and Rapoport, 2007), or financial crises (Bertoli et al., 2010; Yang, 2008a).

[9]This is similar to the argument in Anderson and van Wincoop (2004) that parameter estimates in standard cross-country gravity models of trade flows may suffer from aggregation bias in the presence of unmodeled sector-specific trade costs. Ossa (2012) shows that these biases can explain the small welfare gains from free trade implied by a new class of heterogeneous firm trade models.

tion (see Blundell and Stoker, 2005, 2007). The framework that I propose can be extended to other problems lacking tractable mappings from individual choices to macroeconomic outcomes for which wealth heterogeneity matters. Second, the tests for liquidity constraints are rooted in a permanent income framework (similar to Yang's (2008a) study of return migration to the Philippines), resonate with a rich literature on occupational choice (e.g., Evans and Jovanovic, 1989). Third, the empirical results suggest potential welfare gains from a reallocation of resources across households. If positive income shocks enable previously liquidity-constrained households to send migrants *and* incentivize previously *un*constrained households to retain former migrants, then the ultimate welfare and efficiency gains may be quite large. Additionally, financial barriers to migration may keep surplus labor trapped in rural areas thereby sustaining low agricultural (shadow) wages. Although beyond the scope of the present study, these are important areas for future research in line with the broader literature on resource misallocation and productivity in developing countries (e.g., Hsieh and Klenow, 2009).

The remainder of the paper is organized in five sections. Section 1.2 provides background on migration from Indonesia. The stylized facts presented therein inform the theoretical model in Section 1.3. In Section 1.4, I propose a two-step estimation procedure consistent with the theory and describe the data. Section 1.5 presents empirical results, validation exercises, and estimates of village-specific migration costs. Section 1.6 concludes.

1.2 International Migration from Indonesia

Approximately 700,000 legal contract migrants leave Indonesia annually.[10] The vast majority work abroad for 2-3 years in countries across (South)East Asia and the Middle East. Though Saudi Arabia and Malaysia host the majority of Indonesia's migrant workers, a number of other destinations such as Taiwan and Hong Kong have become increasingly important in recent years. Most migrants work in construction, agriculture, manufacturing, and household services. In recent

[10] Unless otherwise noted, all unreferenced claims in this section are supported by detailed empirical evidence in Bazzi (2012a).

years, women have assumed a growing share of total legal migrant outflows, accounting for 50-80 percent annually. In the remainder of this section, I discuss the relevance of Indonesia for examining financial barriers to international migration.

1.2.1 Migration Flows at the Village Level

This study employs rich data on temporary, legal emigrants from the universe of Indonesian villages in 2005 and 2008. The data come from a triannual administrative census of villages known as Village Potential or by its Indonesian acronym, *Podes*.[11] The data include the total number of village residents working abroad for a "fixed wage and time period." The possible inclusion of undocumented emigrants under this designation does not pose a problem conceptually since these migrants typically face similar decisions and constraints.[12] Moreover, comparisons with other data sources suggest that the Village Potential surveys capture the overwhelming majority of international contract migration flows from Indonesia during this period.

Table 1.1 reveals several stylized facts on international migration from 65,966 Indonesian villages. First, similar to other large developing countries, the number of emigrants is small relative to population size both at the village level and nationally. However, central tendencies can be misleading as labor migration rates are quite high in many rural villages and the figures are scaled by total population rather than working-age population.[13] Second, households in rural areas participate more intensively in international migration than their urban counterparts. Whereas 60 percent of the population resides in rural areas, around 85 percent of migrants hail from these areas. Third, migration rates increase on average by approximately 11 percent between 2005 and 2008. Lastly, the extensive

[11] The data are obtained primarily from key informants in the village government with additional input and corroboration from officials in the subdistrict and district government. Village officials have historically been the first line of bureaucracy from which potential migrants must obtain legal permission to work abroad (Spaan, 1994). Today, these officials authorize the national ID cards required to work outside the country under contract (Bank Indonesia, 2009).

[12] Most undocumented migrants also have similar two to three year contracts (Bank Indonesia, 2009), and the theoretical model does not hinge on any distinction between documented and undocumented migrants.

[13] This is a limitation of the data addressed in robustness checks.

margin cannot be ignored: 45 percent of villages did not have any residents working abroad in 2005, but 40 percent of the national increase in migrant outflows through 2008 originated in these villages. Villages with no migrants differ along several dimensions from villages with migrants recorded in 2005 and/or 2008, an issue to which I return in subsequent sections.

1.2.2 Migration Costs and the Pre-Departure Financial Context

In this subsection, I describe the financial environment facing potential migrants from Indonesia. Although migration costs have fallen over the last decade, these costs still tend to be substantial relative to household income. Important monetary costs include direct placement or recruitment fees. Around 2007, quoted fees ranged from 800 USD for destinations in Asia to 1,200 USD for destinations in the Middle East (Bank Indonesia, 2009). These fees alone approximately equal 75-100 percent of total annual household expenditure in the typical Indonesian household, and 250-400 percent of annual pre-migration wages in Indonesia.[14] Available estimates of total out-of-pocket costs paid prior to departure range from 350 to 900 USD, and around 85 percent of households are unable to finance these costs purely out of own savings (World Bank, 2009, 2010).

Despite opportunities for financial and legal innovation in this environment, the credit market for potential international migrants has been thin. Most formal lenders view migrants as high risk borrowers given the difficulty of tendering cross-border repayments and the lack of creditworthiness of potential family co-signers. Less than 5 percent of migrants report borrowing from formal financial institutions, and an estimated 80 percent of upfront costs are financed through informal borrowing from friends, family, and recruitment agencies (World Bank, 2010).

Although some recruiters offer interlinked contracts that allow migrants to borrow against future earnings, aspiring migrants are still required to pay upfront some fraction of total pre-departure and placement costs prior to earning the first

[14](i) is based on nationally representative survey data (*Susenas*) in 2007, and (ii) is based on survey data in Bank Indonesia (2009).

month's wage. These "downpayments" often serve as a commitment device for recruiters soliciting new migrants. There also tends to be variation across destinations and genders in the magnitude of upfront costs, and in the absence of local (and financially liquid) recruiters, individuals must either pay the full migration costs upfront or the costs of identifying and reaching more capable recruiters in nearby urban centers. In practice, in rural villages with no recent migration history, first-time international migration is nearly impossible without recruiters.

Furthermore, even those interlinked contracts requiring little cash-in-advance have potentially large implicit financial barriers. Such debt contracts often impart effective annual interest rates over 60 percent (World Bank, 2010) and withhold partial or entire earnings for many months after beginning employment in the destination country.[15] Coupled with 1-3 months of pre-departure training without pay, these financing requirements may constrain migration choice in households unable to cope with an extended period of lost income by a productive member. If households require returns to investment (i.e., remittances) earlier than allowable under the debt contract, then financial constraints could prove binding even in the unlikely case where an interlinking arrangement requires no pre-departure financing by the migrant.[16] Nevertheless, insomuch as recruitment agencies ease the predeparture liquidity constraints facing potential migrants, their presence throughout Indonesia should work against finding prevalent financial barriers to migration.

In sum, although informal solutions exist in some circumstances, determining whether financial barriers actually constrain international migration flows requires a new theoretical framework relating changes in migration rates to exogenous income shocks, the empirical context for which I discuss next.

[15]Despite legal stipulations mandating no more than 20 percent of monthly wages be withheld, typical deductions are around 100 percent for the first 6 months in Malaysia, 75 percent for the first 5-6 months in Singapore, 80 percent for the first 7 months in Hong Kong, 67 percent for the first 15 months in Taiwan; and 100 percent for the first 3 months in Saudi Arabia (MICRA, 2008).

[16]In a related context, Field et al. (2011) show how microfinance debt structuring can reduce investment in profitable microenterprises.

1.2.3 Why Rainfall and Rice Prices Matter for Migration Decisions

I highlight here several reasons why rainfall and rice price shocks should capture changes in incomes relevant to international migration choice in Indonesia. First, in addition to the fact that nearly 90 percent of migrants hail from rural areas, the majority come from primarily agricultural households. Second, women, who comprise the majority of (legal) labor migrants, account for 40-45 percent of total agricultural labor employed in rice cultivation. Third, international migration tends to be countercyclical with respect to the rice planting season. Migrant outflows measured at the district level (in auxiliary administrative data) are 10-12 percent lower during months falling within the local growing seasons.

Fourth, based on nationally representative household survey data from 2005, Figure 1.2 reveals that households with migrants tend to be drawn from the middle of the distribution of agricultural landholdings. This cross-sectional inverted U relationship is consistent with evidence from other developing countries (e.g., see McKenzie and Rapoport, 2007, on Mexico). One interpretation of the graph is that the land-poor cannot afford to migrate while the land-rich lack the incentives to do so. The theory developed below admits this possibility and treats landholdings as the key source of observable heterogeneity through which rainfall and rice price shocks affect income and ultimately the decision to migrate.

As in many developing countries, landholdings are a fundamental means of income generation for the majority of Indonesians in rural areas. According to the Agricultural Census in 2003 (see Section 1.4.2), roughly 22.3 million out of 39.6 million households own or rent some agricultural land. Among the landholding population, 54 percent of households control some wetland (or *sawah*) particularly germane to rice production, and 58 percent of households report growing rice in the 2002-3 growing season. An estimated 70-75 percent of rice-growing households are net rice producers while 30 percent of all farm households were net rice producers as of 2003 (McCulloch, 2008). In practice, land is often the most valuable asset under the control of relatively poor households, but it is also quite illiquid given the thin or missing land markets prevailing in most of rural Indonesia. Ultimately, landholdings could serve several purposes relevant to temporary migration decisions. For tractability, the theoretical model developed in the next section focuses on the primary role of land in generating income.

1.3 Theoretical Model

This section develops a microfounded model of temporary international migration flows. At the time of decision making, individuals face uncertainty over future agricultural income at home while learning about potential net wages from working abroad next period.[17] However, liquidity constraints prevent some households from making otherwise standard migration decisions based on comparing expected net income at home and abroad. The migration decision is therefore couched as a short-term investment opportunity with no recourse to formal, collateralizable financing.[18]

As the main source of income heterogeneity, agricultural landholdings alleviate financial constraints to migration and incentivize further allocation of household labor to domestic rice production. The tradeoff therein distinguishes this model from Roy-type models of migration in which the main sources of heterogeneity, observable human capital and unobserved skill, are portable across international borders.[19] By exploiting the Pareto distribution for landholdings and adopting modeling insights from heterogeneous firm trade theory (Melitz, 2003), I am able to map the welfare effects of agricultural income shocks into tractable expressions for migration flows measured at the village level.

[17]For the results of the model to hold, we merely require that individuals within the village have identical expectations over net incomes on offer regardless of the underlying sources or degree of uncertainty.

[18]The formulation of liquidity constraints is therefore broad enough to encompass the opportunity cost of time foregone working without wages implicit in the possible interlinked contracts mentioned above. Following a standard definition of liquidity constraints (Hayashi, 1987): potential migrants can be deemed liquidity constrained if they face quantity constraints in the amount of borrowing or if loan terms available to them as borrowers are less favorable than those at which they could lend.

[19]Landholdings are (pre)determined exogenously and cannot be liquidated to finance migration costs. While inappropriate over the long-run, this assumption is reasonable when studying temporary migration from a developing country with missing land markets.

1.3.1 Model Environment

Suppose home income for individual $i = 1, \ldots, N_v$ in village $v = 1, \ldots, V$ in period t is given by $\Pi_{ivt} = p_{vt} Y_{ivt}$, where p_{vt} is the exogenously given farmgate price (net of distribution costs) for one unit of commodity $\mathcal{Y}$ (say, rice) produced in village v. Agricultural output is produced according to a constant returns to scale technology: $Y_{ivt} = \sigma_{vt} K_v^{\theta} S_{iv}^{\phi} R_{iv}^{\beta}$ where σ_{vt} is the level of rainfall, K_v is publicly available capital, S_{iv} is individual i's efficiency units of labor, and R_{iv} is i's household landholdings in hectares.

Individual migrant i from village v can earn gross wages W_{ivjt} net of costs C_{vjt} abroad in j in period t.[20] Gross wages W_{ivjt} are a function of skill (in efficiency units) S_{iv}, a cultural discount factor δ_{vj}, and a time-varying demand term d_{jt} common across villages. Migration costs C_{vjt} are fixed across individuals but vary across villages and time. Conditional on these terms, all Indonesian nationals face the same wage in destination j. At the time of decision-making at the end of $t-1$, individuals learn net wages as stipulated in contracts offered by recruiters or local network intermediaries.

Each period, individual i earns land income, then allocates this between consumption and financing future migration, with no other savings. Given data constraints, I abstract away from intra-household issues, treating individual choice as tantamount to that of a collective household.[21] In an unconstrained setting, the collective household sends family member i abroad next period if her net returns to migration exceed the foregone expected income (or marginal revenue product of her labor) at home[22]

$$W_{ivj,t+1} - C_{vj,t+1} \geq \mathbb{E}_t[p_{v,t+1}\sigma_{v,t+1}] K_v^{\theta} S_{iv}^{\phi} R_{iv}^{\beta} \ . \tag{1.1}$$

[20]In effect, j is a composite destination since empirically I only observe total village migration across all destinations.

[21]This turns out to be a conservative approach. In a model with similar primitive conditions, Delpierre (2012) shows that introducing intra-household bargaining—in particular, allowing for imperfect commitment (to remit) between the migrant and remaining members—tends to exacerbate the financial barriers to investment in profitable migration opportunities.

[22]Hence, there is no tradeoff between holding on to one's land and migrating as in Jayachandran (2006).

However, only those individuals with enough savings from the prior period can afford to cover the portion $\tau_{vj} \in [0,1]$ of the costs $C_{vj,t+1}$ of migration to destination j that must be paid prior to earning the first month's wage abroad. The τ_{vj} parameter operationalizes liquidity constraints in the model. If $\tau_{vj} = 0$, then equation (1.1) suffices to identify individual i's migration choice. If $\tau_{vj} > 0$, the financing arrangements available to individuals in village v—either through (in)formal credit markets or interlinked contracts offered by recruiters—will not cover all predeparture migration costs (including foregone time without wages). In this case, the fixed migration cost imposes a minimum wealth requirement R_L in order to migrate next period. Combining both conditions, individuals with the following landholdings will be abroad in $t+1$

$$\underbrace{\left(\frac{\tau_{vj} C_{vj,t+1}}{p_{vt}\sigma_{vt} K_v^{\theta} S_{iv}^{\phi}}\right)^{\frac{1}{\beta}}}_{R_L} \leq R_{iv} \leq \underbrace{\left(\frac{W_{ivj,t+1} - C_{vj,t+1}}{\mathbb{E}_t[p_{v,t+1}\sigma_{v,t+1}] K_v^{\theta} S_{iv}^{\phi}}\right)^{\frac{1}{\beta}}}_{R_U}. \tag{1.2}$$

Note that this expression preserves the (stark) inverted U relationship between landholdings and migration choice found in the related model of McKenzie and Rapoport (2007) and observed above in Indonesian data.

Inequality (1.2) contains the key cross-sectional relationship between prices, rainfall, and migration choice. However, in order to identify the relationship between changes in individual income and village-level migration rates, we must make some distributional assumptions. I now develop these assumptions in a manner that is guided by the data and discuss robustness along the way.

Distributional Assumptions

Landholdings (R_{iv})[23] For both empirical and theoretical reasons, I approximate the landholdings distribution with village-specific Pareto distributions. Empirically, the distribution of landholdings in Indonesia is well represented by a power law. Unlike other empirical phenomena (e.g., income, city size, firm pro-

[23]See Appendix 1.9.6 for supporting evidence and further discussion pursuant to the landholdings distribution.

ductivity; see Gabaix, 2009), however, the familiar log-linearity of the complementary cumulative distribution function (CCDF) is not restricted to the very upper tail. This key property becomes increasingly apparent for landholdings $R_{iv} > 0.1$ hectares (Ha) (around the 15th percentile nationally). Figure 1.3 demonstrates this using log rank–log size plots for 16 Indonesian district. The approximate linearity in these graphs holds for administrative divisions down to the village-level.

Analogous to the logic underlying Zipf's law for cities (Gabaix, 1999), the Pareto landholdings distribution could arise in the steady state from a random population growth process on a fixed land mass. Allowing the Pareto dispersion parameter to differ across villages provides a convenient way to capture spatial variation in this long-run process. Increasing urbanization, the demographic transition, and the structural transformation meanwhile may exert precisely the sort of stabilizing pressure giving rise to Pareto properties (i.e., counteracting forces towards exponentiality that may have prevailed in earlier periods driven by inheritance, higher birth rates, and more limited rural to urban migration).

Formally, within each village v, there is a continuum of individuals i with landholdings R_{iv} drawn from a Pareto distribution with village-specific dispersion parameter λ_v above some minimum threshold $\underline{R}$ common across villages. The density function is given by $\lambda_v \underline{R}^{\lambda_v} R_{iv}^{-\lambda_v - 1}$ where $\lambda_v > 0$.[24] The mean and variance of landholdings are decreasing in λ_v, which provides a sufficient statistic for landholdings inequality (the familiar Gini coefficient $G_v = 1/(2\lambda_v - 1)$). A nice feature of the Pareto distribution is that its shape is preserved over all truncated segments of the distribution above $\underline{R}$ and hence is invariant to the location of the land wealth thresholds, R_L and R_U, for migration in inequality (1.2). The Pareto formulation allows these thresholds to differ across villages depending on prevailing economic conditions. This is not possible in an important related model in Jayachandran (2006), which treats landholdings as binary. In Section 1.4.2, I discuss the practical implications of the Pareto assumptions for the empirical analysis.

Efficiency units of labor (S_{iv}) Skill heterogeneity is common in models of mi-

[24]Although an upper truncation $\overline{R}$ is arguably a realistic feature of land availability, an infinite upper bound simplifies the notation and calculations considerably without compromising the key features of the theoretical model.

gration choice. To allow incomes to be increasing in education levels, I assume individuals have high skill S^H with probability γ_v and low skill S^L with probability $1-\gamma_v$ such that $S_{iv} = I \times S^L + (1-I) \times S^H$ and $R_{iv} \perp I \sim$ bernoulli(γ_v). The bernoulli formulation is chosen for simplicity and is without loss of generality, and $R_{iv} \perp I$ ensures that landholdings remain the fundamental source of idiosyncratic productivity at home. The key model predictions are robust to relaxing this assumption or to ignoring skill heterogeneity altogether.[25]

Producer prices (p_{vt}) and rainfall (σ_{vt}) Producer prices follow an $ARMA(1,q)$ process with possibly heterogeneous AR parameters, $p_{vt} = \alpha_v p_{v,t-1} + \sum_{s=0}^{q} \theta_s e_{v,t-s}$ where $\theta_0 = 1$ and e_{vt} is a mean-zero shock. This specification is sufficiently general to encompass unit root processes ($\alpha_v = 1$) and the associated permanent effect of price shocks. Meanwhile, taking a standard approach in the literature (see Rosenzweig and Wolpin, 2000), rainfall follows an i.i.d. process such that $\sigma_{vt} = \overline{\sigma}_v + a_{vt}$, where $\overline{\sigma}_v$ is the long-run average level of rainfall in village v and a_{vt} is a mean-zero shock in year or growing season t.[26]

Migration costs (C_{vjt}) and upfront cost share (τ_{vj}) The cost of migration from village v is comprised of (i) observable components that are a function of distance to legal emigrant processing hubs, attractiveness to recruiter agencies, and general remoteness, and (ii) time-varying components that are observable to village residents but unobservable to the researcher. The key assumption here is that individuals within the same local area face identical fixed upfront costs of migration—regardless of whether those costs are paid to outside recruiters or members of the village social network.[27]

[25]Munshi (2003) makes an assumption similar to $R_{iv} \perp I$ in his model. Assuming instead that R_{iv} is drawn jointly with S_{iv} would leave the main qualitative predictions of the model mostly unchanged so long as S_{iv} and R_{iv} are positively correlated and the elasticity of foreign wages with respect to R_{iv} is not too large. This latter assumption is reasonable given that wages exhibit very little variation within occupation×destination and rural Indonesians largely work in a few low-skill occupations (see Bank Indonesia, 2009).

[26]In Appendix 1.9.5, I show that these formulations for prices and rainfall are consistent with Indonesian data.

[27]This seems reasonable given (i) the small size of most Indonesian villages (see Table 1.1), and (ii) the explosion in recruitment activity over the last decade (see Bachtiar, 2011), which has increased pressures towards competitive upfront cost-pricing. If, however, local recruiters or network intermediaries can impose higher upfront costs on wealthier households, then τ_{vj} might be correlated with R_{iv}. In this case, the main qualitative predictions of the model remain

1.3.2 Identifying the Existence and Magnitude of Liquidity Constraints

In characterizing migration flows, I adopt the following timing convention. Migration rates observed at the beginning of periods t and $t+1$ are the outcome of collective household decision-making at the end of prior periods $t-1$ and t, respectively, at which time the net wage schedule on offer next period is presented to potential migrants. Additionally, I (i) impose the innocuous normalization $\underline{R} = 1$ Ha and (ii) have implicitly integrated over S_{iv} when referring to R_L and R_U and using ω_{vjt} to denote the prevailing gross foreign wage offer for village v. I refer to the following four equations as the *intensive margin*.

Suppose first that liquidity constraints are binding for some households in village v so that $\tau_{vj} > 0$ and $R_L \geq \underline{R}$. The earlier distributional assumptions imply the following *flow* migration rate between periods

$$\Delta\left(\frac{M_{v,t+1}}{N_{v,t+1}}\right) = \left(K_v^{\theta} S_v^{\phi}\right)^{\frac{\lambda_v}{\beta}} \Delta\left\{p_{vt}^{\frac{\lambda_v}{\beta}}\left[\left(\frac{\overline{\sigma}_v + a_{vt}}{\tau_{vj} C_{vj,t+1}}\right)^{\frac{\lambda_v}{\beta}} - \left(\frac{\alpha_v \overline{\sigma}_v}{\omega_{vj,t+1} - C_{vj,t+1}}\right)^{\frac{\lambda_v}{\beta}}\right]\right\}, \quad (1.3)$$

where N_{vs} is village v population in period s and M_{vs} is the number of residents working abroad in s. Expressing the flow migration rate in log rather than level differences gives

$$\Delta \ln\left(\frac{M_{v,t+1}}{N_{v,t+1}}\right) = \frac{\lambda_v}{\beta} \Delta \ln p_{vt} + \Delta \ln\left[\left(\frac{\overline{\sigma}_v + a_{vt}}{\tau_{vj} C_{vj,t+1}}\right)^{\frac{\lambda_v}{\beta}} - \left(\frac{\overline{\sigma}_v \alpha_v}{\omega_{vj,t+1} - C_{vj,t+1}}\right)^{\frac{\lambda_v}{\beta}}\right]. \quad (1.4)$$

If, on the other hand, $R_L < \underline{R}$, then liquidity constraints are not binding among potential migrants in village v. (Note that although $\tau_{vj} = 0 \implies R_L < \underline{R}$, the reverse implication need not hold.) In this case,

$$\Delta\left(\frac{M_{v,t+1}}{N_{v,t+1}}\right) = \left(\alpha_v \overline{\sigma}_v K_v^{\theta} S_v^{\phi}\right)^{\frac{\lambda_v}{\beta}} \left[\left(\frac{p_{v,t-1}}{\omega_{vjt} - C_{vjt}}\right)^{\frac{\lambda_v}{\beta}} - \left(\frac{p_{vt}}{\omega_{vj,t+1} - C_{vj,t+1}}\right)^{\frac{\lambda_v}{\beta}}\right]. \quad (1.5)$$

Taking log differences meanwhile implies

$$\Delta \ln\left(\frac{M_{v,t+1}}{N_{v,t+1}}\right) = \ln\left[1 - \left(\frac{\alpha_v p_{vt} \overline{\sigma}_v K_v^{\theta} S_v^{\phi}}{\omega_{vj,t+1} - C_{vj,t+1}}\right)^{\frac{\lambda_v}{\beta}}\right] - \ln\left[1 - \left(\frac{\alpha_v p_{v,t-1} \overline{\sigma}_v K_v^{\theta} S_v^{\phi}}{\omega_{vjt} - C_{vjt}}\right)^{\frac{\lambda_v}{\beta}}\right] \quad (1.6)$$

---

unchanged, but the magnitude of migration flows might differ.

Note that in the presence of liquidity constraints, log-linearization removes the time-invariant determinants of agricultural output $(K_v^{\theta} S_v^{\phi})$ and provides a convenient linear expression relating the price shock to changes in migration rates. Log-linearization also highlights the fact that price *shocks* matter in the presence of liquidity constraints whereas price *levels* matter in their absence.

The expressions above give rise to multiple testable implications. In Propositions 1 and 2 below, I focus on those aspects of the model that make it possible to identify the presence and magnitude of liquidity constraints using available data. The propositions are based on the log-linearized expressions in (1.4) and (1.6), and the proofs can be found in Appendix 1.9.1.

Proposition 1 *If liquidity constraints are not binding for any households in village v, then the flow migration rate $\Delta \ln (M_{v,t+1}/N_{v,t+1})$ between periods t and $t+1$ is uncorrelated with rainfall shocks a_{vt} and $a_{v,t-1}$, and decreasing in recent prices p_{vt} (and increasing in distant prices $p_{v,t-1}$ via the negative effect on $\ln M_{vt}/N_{vt}$). Conversely, if liquidity constraints are binding, then the flow migration rate is (i) increasing in recent rainfall shocks a_{vt} (and decreasing in distant rainfall shocks $a_{v,t-1}$ via the positive effect on $\ln M_{vt}/N_{vt}$), and (ii) increasing in price shocks $\Delta \ln p_{vt}$.*

This proposition delivers a simple empirical test for the existence of liquidity constraints based on the sign of two coefficients. Consistent with the literature on the permanent income hypothesis, household migration choice should only be affected by transitory shocks if liquidity constraints are binding. Potentially permanent shocks should affect migration choices in either case. However, these predictions only identify the average effect of agricultural income shocks and do not tell us anything about whether and how this effect varies across villages depending on, for example, the distribution of agricultural productivity.

Proposition 2 *(i) In the presence of liquidity constraints, price shocks $\Delta \ln p_{vt}$ have larger positive effects on the flow migration rate in villages with a greater mass of small landholders (high λ_v). In the absence of liquidity constraints, increases in recent prices p_{vt} have larger negative effects on the flow migration rate in villages with a greater mass of large landholders (low λ_v). (ii) In the presence* *of liquidity constraints, recent rainfall shocks a_{vt} have larger positive effects on the flow migration rate in villages with a greater mass of small landholders.*

This proposition reveals how the distribution of wealth affects the extent to which income shocks increase or decrease migration flows. If liquidity constraints are binding in the population of potential international migrants, then income shocks should have the strongest effects on migration choice among the poor. Thus, all else equal, positive rainfall and price shocks should induce greater migration flows from villages with a relatively higher share of poor households (i.e., villages with a greater mass of small landholders). Deaton (1991) shows that positive serial correlation in the income process, as in the $ARMA$ formulation for prices, reduces the scope for income smoothing among liquidity-constrained households. A large positive income shock relaxing some of those constraints might then also make it possible for poor households to undertake novel risk diversification measures such as international migration. Moreover, to the extent that positive income shocks loosen informal credit markets, we should expect a larger migration response in villages with less *ex ante* inequality where the scope for inter-household borrowing was more limited. Assuming no sources of external finance, the dispersion parameter λ_v captures the *potential* thickness of these informal credit markets in the village. Applying a structural interpretation to the model, the cross-partial effect of price shocks and λ_v on the flow migration rate is exactly equal to $1/\beta$ (i.e., the inverse of the share of land in the production function).

On the other hand, if cash-in-advance constraints are not binding in the population of potential migrants, then rising output prices leads to a fall in migration flows on account of rising income at home. Proposition 2 suggests that this reduction in migration flows should be steeper in villages with a greater mass of large landholders. This differential reduction occurs primarily because price increases provide a stronger disincentive to migrate (via higher expected future prices) among higher landholding households but also perhaps because price increases may lead to a loosening of credit markets thereby allowing low landholding households to borrow out of the increased income of their wealthier neighbors to finance further investment in agriculture.[28]

1.3.3 A Theoretical Characterization of the Extensive Margin

There are several reasons why the barriers to migration may differ between villages with and without migrants (e.g., unobservable networks, Carrington et al., 1996). Unlike existing random utility models of migration, the theoretical model outlined above makes it possible to characterize these differences in the probability of (non-)zero migration or what I refer to as the *extensive margin.* Although equations (1.3)-(1.6) implicitly assume non-zero migrant stocks in both periods, identically zero migration can arise as an equilibrium outcome under certain conditions. Taking the theory to data requires formalizing these conditions. If $\tau_{vj} > 0$ and $R_L \geq \underline{R}$, the *stock* migration rate (M_{vt}/N_{vt}) equals zero whenever the maximum village v landholding $\max_k R_{kv} \equiv \widetilde{R}_v < R_L$ or the minimum landholding $\min_\ell R_{\ell v} \equiv \underset{\sim}{R}_v > R_U$. If $R_L < \underline{R}$ $(\Longleftarrow \tau_{vj} = 0)$, $M_{vt}/N_{vt} = 0$ whenever $\underset{\sim}{R}_v > R_U$.[29] To observe any migrants from village v, then, at least one individual must be able to afford to migrate and at least one individual must deem migration profitable. Focusing on the case where $\tau_{vj} > 0$ and $R_L \geq \underline{R}$ and noting that $\underset{\sim}{R}_v$ and $\widetilde{R}_v$ are order statistics drawn from N_v *i.i.d.* Pareto random variables, the law of total probability implies the following extensive margin:

$$\mathbb{P}(\underset{\sim}{R}_v \leq R_U,\ \widetilde{R}_v \geq R_L) = 1 - R_U^{-\lambda_v N_v} - \left(1 - R_L^{-\lambda_v}\right)^{N_v}, \tag{1.7}$$

where $\widetilde{R}_v$ follows a known Stoppa distribution (see Kleiber and Kotz, 2003) and $\underset{\sim}{R}_v$ follows a Pareto distribution with shape parameter $\lambda_v N_v$. This finite sample formulation can be rationalized by appealing to the truncation in equation (1.2)

[28] It is worth noting that the key qualitative predictions of the model are robust to allowing for idiosyncratic preferences (or disamenity costs) over living at home and abroad. As in standard migration choice models, I can assume that those preferences follow an exponential (as in Klein and Ventura, 2009) or extreme value distribution (as in Grogger and Hanson, 2011). So long as preferences are orthogonal to landholdings, Propositions 1 and 2 hold. However, this generalization comes at the expense of concise log-linearized expressions for migration flows.

[29] Note that the common truncation $\underline{R}$ across villages ensures that zero migration cannot be derived in an *ad hoc* manner.

and the practical limits of village population size.[30] In this setting, heavily populated villages are relatively more likely to have any international migrants. This relationship has both statistical and economic content as I discuss in the next section. Taking a population approach, however, we let $N_v \to \infty$, in which case $\widetilde{R}_v \sim$ Fréchet(λ_v) and $\underset{\sim}{R}_v \sim$ Weibull$(1, \lambda_v)$ in the limit (see Gumbel, 1958).[31] Since $\widetilde{R}_v \perp \underset{\sim}{R}_v$ asymptotically, equation (1.7) becomes:

$$\mathbb{P}(\underset{\sim}{R}_v \leq R_U,\ \widetilde{R}_v \geq R_L) = \left(1 - e^{-R_U^{\lambda_v}}\right) \times \left(1 - e^{-R_L^{-\lambda_v}}\right). \tag{1.8}$$

Zero migration can still arise in this case but does so irrespective of the number of potential migrants. When $R_L < \underline{R}$, equations (1.7) and (1.8) simplify, respectively, to $1 - R_U^{-\lambda_v N_v}$ and $1 - e^{-R_U^{\lambda_v}}$.

Equations (1.7) and (1.8) imply an ambiguous relationship between the distribution of landholdings and the extensive margin when $\tau_{vj} > 0$ and $R_L \geq \underline{R}$. For given R_L and R_U, the probability that any village residents find migration profitable is increasing in λ_v whereas the probability that any residents can afford to migrate is decreasing in λ_v. This ambiguity differs from that along the intensive margin implied by Proposition 2, in that for zero migration villages, we have two cases: (i) all households cannot afford to send migrants despite available income gains, or (ii) all households can afford to send migrants but the relative income gains are insufficient. The empirical analysis of the extensive margin will be informative as to which threshold "matters more" on average: the probability of non-zero migration is decreasing in λ_v if case (i) prevails for the average village and increasing in λ_v if case (ii) prevails. However, if $R_L < \underline{R}$, then case (ii) should hold regardless.

[30]Eaton et al. (forthcoming) apply a similar rationale in a gravity model with a finite number of heterogeneous firms.

[31]This approach is analogous to that in the Helpman et al. (2008) gravity model with a continuum of heterogeneous firms.

1.4 Empirical Strategy

In this section, I develop the empirical strategy for evaluating the theory. I begin by arguing in favor of a two-step estimating framework, which has theoretical and practical advantages over existing approaches in the migration literature. Then, I propose candidate exclusion restrictions to identify second-step parameters for the intensive margin. Lastly, I describe data and measurement of key variables in the model.

To begin, note that the model implies the following expected stock migration rate in period s:

$$\mathbb{E}\left(\frac{M_{vs}}{N_{vs}}\right) = \underbrace{\mathbb{E}\Big(\mathbf{1}\{R_{Ls} \leq R_{iv} \leq R_{Us}\} \;\Big|\; \widetilde{R}_v \geq R_{Ls}, \underset{\sim}{R}_v \leq R_{Us}\Big)}_{\text{intensive margin: eq. (1.3)}} \times \underbrace{\mathbb{P}\Big(\widetilde{R}_v \geq R_{Ls}, \underset{\sim}{R}_v \leq R_{Us}\Big)}_{\text{extensive margin: eq. (1.7)}}. \tag{1.9}$$

Thus, any estimated relationship between agricultural income shocks and the unconditional migration rate $\left(\frac{M_{vs}}{N_{vs}}\right)$ (i.e., including zeros) will reflect a mixture of the extensive and intensive margin distributions. The log-linearization suggested by equations (1.4) and (1.6) implies taking logs of the migration rate and hence focusing only on the intensive margin. However, doing so ignores the important observable and unobservable differences between villages above and below the extensive margin thresholds. Put simply, the shocks generating transitions from zero to one migrant could be quite different from those generating movement along the intensive margin. In order to account for these differences, I propose a two-period Heckman (1976) approach as an alternative to existing strategies for handling zeros in the migration literature.[32]

There are a few key reasons for favoring a two-step approach to estimating the extensive and intensive margins of international migration flows. By failing to account explicitly for the entry of rural villages into international labor markets, existing estimation strategies make it difficult to isolate where and to what

[32] These include OLS on the migration rate (Mayda, 2010) or the log migrant stock +1 (Ortega and Peri, 2009), standard Tobit (Mayda, 2010), threshold Tobit (Simpson and Sparber, 2010), restricting to non-zero observations (Grogger and Hanson, 2011), and Poisson QMLE (Beine et al., 2011). In practice, the Heckman (1976) two-step approach here nests alternative two-part models, which do not explicitly account for the influence of the first-step estimates on the second-step estimates (see Leung and Yu, 1996).

extent financial—or informational or policy—constraints are binding. Intuitively, the barriers constraining outflows from villages with a long history of international migration plausibly differ from those in villages with no recent connection to international labor markets. For example, in villages with no migrants, informational constraints may be relatively more binding than financial constraints in preventing emigration among the poor whereas the opposite may be true in villages with a long migration history. The theory in Section 1.3 provides justification for an econometric method allowing for such differences,[33] and a large literature on networks and the persistence of migration flows supports the underlying intuition (e.g., Munshi, 2003).

Beyond the theoretical benefits of the two-step approach, there is also an important practical advantage to taking logs of the migration rates. The distribution of stock migration rates M_{vs}/N_{vs} is characterized not only by a preponderance of zeros but is also heavily right skewed above zero. Figure 1.5 demonstrates this feature of the data. The bottom panel of Figure 1.5 then shows how the log transformation provides a more readily interpretable flow migration rate relative to the specification in levels with or without zeros. This has important implications for how we model the error term.

The latent variable estimation strategy that I propose has parallels in the estimation of (i) the labor supply elasticity in the presence of non-participation (Blundell et al., 2011), and (ii) demand system parameters in the presence of zero consumption (Yen, 2005). Specifically, the suggested setup focuses on those villages with positive migration rates while also implicitly considering what the *potential* nonzero migration rates would be if villages with zero migration subsequently entered international labor markets.

In the general two-step procedure, I estimate the log flow migration rate between periods t and $t+1$ conditional on first-stage equations for the extensive

[33] Moreover, I show in Appendix 1.9.7 that the empirical incidence of zero migration is not a statistical artifact. Adapting a simple test developed in the trade literature (Armenter and Koren, 2010), I compare the empirical incidence of zeros with that arising from a model in which villages (bins) receive migrants (balls) randomly but with probability proportional to village population. The incidence of zeros in the data is much higher than would be predicted on the basis of this random balls-and-bins allocation. According to the test, only 5.5 percent of the 27,297 zeros in the 2005 data can be deemed an atheoretical regularity in sparse data.

margin in each period. The three-equation system accounts for entering ($M_{vt} = 0, M_{v,t+1} > 0$) and exiting villages ($M_{vt} > 0, M_{v,t+1} = 0$) (around 20 percent of villages) and also allows the unobservable extensive margin thresholds to be correlated across periods:

$$\begin{aligned} m^*_{vt} &= \eta'_{t-1}\mathbf{Z}_{v,t-1} + u_{vt}; \quad m^*_{v,t+1} = \eta'_t\mathbf{Z}_{vt} + u_{v,t+1}, \\ \Delta \ln \left(\frac{M_{v,t+1}}{N_{v,t+1}}\right) &= \boldsymbol{\zeta}'\Delta\mathbf{X}_{vt} + \Delta\varepsilon_{v,t+1} \qquad \text{iff} \quad m^*_{vt} > 0 \text{ and } m^*_{v,t+1} > 0, \end{aligned} \tag{1.10}$$

where m^*_{vs} is a continuous latent variable, and $\mathbf{Z}_{vs}$ and $\mathbf{X}_{vs}$ comprise, respectively, the determinants of the extensive and intensive margin. This setup could alternatively be construed as a panel sample selection problem by including village fixed effects in the latent variable equations (see Kyriazidou, 1997; Rochina-Barrachina, 1999; Wooldridge, 1995).[34] Instead, I propose a linear parametrization of the village fixed effects that controls for some of the important time-invariant determinants of migration suggested by the model. This allows us to retain the rich information content of always- ($m^*_{vs} > 0 \ \forall s$) and never-migrant ($m^*_{vs} \leq 0 \ \forall s$) villages while also identifying differential effects of fixed village characteristics on the extensive margin across periods. Then, to account for the additional term in the conditional expectation,

$$\mathbb{E}\left[\Delta \ln \left(\frac{M_{v,t+1}}{N_{v,t+1}}\right)\right] = \boldsymbol{\zeta}'\Delta\mathbf{X}_{vt} + \mathbb{E}\left[\Delta\varepsilon_{v,t+1}\middle|\eta'_{t-1}\mathbf{Z}_{v,t-1} > -u_{vt}, \eta'_t\mathbf{Z}_{vt} > -u_{v,t+1}\right],$$

I employ parametric and semiparametric correction procedures. Although tractable, the parametric approach originally due to Poirier (1980) has strong distributional assumptions. As an alternative, I use a variation on a semiparametric approach due to Das et al. (2003) that includes in the second-stage a flexible function of first-stage propensity scores. Appendix 1.9.2 details both procedures.

The setup in (1.10) comes out of a latent variable framework suggested by

[34] With two additional periods, I could employ the more flexible dynamic panel sample selection approach of Gayle and Viauroux (2007).

the theory. Note that

$$\frac{M_{v,t+1}}{N_{v,t+1}} = \left(R_{Lt}^{-\lambda_v} - R_{Ut}^{-\lambda_v}\right) \times \Lambda_{vt},$$

where $\Lambda_{vt} = \mathbf{1}(\widetilde{R}_v \geq R_{Lt}) \times \mathbf{1}(\underset{\sim}{R}_v \leq R_{Ut})$. Two latent variables can thus be defined in terms of $\underset{\sim}{R}_v$, $\widetilde{R}_v$, and unobservable village-level migration costs implicit in C_{vjt}:

$$Z_{vt}^{l} = \frac{p_{vt}\sigma_{vt}K_v^{\theta}S_v^{\phi}\widetilde{R}_v^{\beta}}{\tau_{vj}C_{vj,t+1}}; \qquad Z_{vt}^{w} = \frac{\omega_{vj,t+1} - C_{vj,t+1}}{\mathbb{E}_t[p_{v,t+1}\sigma_{v,t+1}]K_v^{\theta}S_v^{\phi}\underset{\sim}{R}_v^{\beta}} \quad . \tag{1.11}$$

The equation for m_{vt}^{*} is then a compact expression for the composite latent variable comprised of Z_{vt}^{l} and Z_{vt}^{w}.

The latent variable formulation proves useful both theoretically and empirically. If one considers $\underset{\sim}{R}_v$ and $\widetilde{R}_v$ as unobservable, then equations (1.7) and (1.8) provide a convenient way to relate a single parameter of the landholdings distribution (λ_v) in village v to the extensive margin. On the other hand, with universal agricultural census data, I do, in fact, observe the actual landholdings extrema for every village v. In this case, equations (1.11) suggest that the probability of having any migrants is increasing in the log difference between the maximum and minimum landholding sizes in village v, $\ln \widetilde{R}_v - \ln \underset{\sim}{R}_v$. However, unlike λ_v, the extreme order statistics do not directly affect the intensive margin.

1.4.1 Exclusion Restrictions

The key assumption of the estimating framework in (1.10) is that the error in the flow migration rate equation is a multiple of the errors in the extensive margin, plus some noise independent of the extensive margin. This seems reasonable given the theoretical structure around the two margins. However, credible identification of the second stage parameters $\boldsymbol{\zeta}$ requires that a subset of variables in $\mathbf{Z}$ shift the extensive margin for village v while not affecting the intensity with which its residents participate in international labor markets.[35] While the theory

[35] Although exclusion restrictions are theoretically unnecessary in the parametric model (Wilde, 2000), their use strengthens the case for model robustness and is moreover required for identification in the semiparametric model. See Appendix 1.9.2.

does not impose exact exclusion restrictions, certain features of the model and the empirical context give rise to a set of candidate instruments.

First, consider (log) maximum and minimum landholdings within the village. Intuitively, the range of landholding sizes is informative about the poorest and wealthiest among the population of potential migrants within the village. Whether the wealthiest finds migration affordable and the poorest finds migration profitable are sufficient to identify nonzero migration. However, neither are informative about the share of the population that finds migration profitable *and* affordable. Hence, both are plausibly excludable.

The finite sample formulation for landholdings extrema gives rise to another potential exclusion restriction: village population size. In terms of equation (1.7), the expected location of the maximum (minimum) landholdings is increasing (decreasing) in village population. In this respect, population size demarcates the boundaries of potential wealth and informal credit markets within the village. Additionally, the population size instrument purges the (minimal) purely statistical and atheoretical variation in the extensive margin (see footnote 33). Another way in which population size affects the extensive margin is through potential migrant market size as perceived by recruiters based in cities.

A simple yet realistic framework for recruiter location choice generates additional candidate instruments. In theory, τ_{vj} and C_{vjt} internalize the foreign demand for migrants in destination j as well as the market potential in and cost of serving village v. However, if the market for potential migrants is too small, recruiters will not serve village v, and $\tau_{vj}C_{vjt}$ will be prohibitively high in some villages. Given the difficulty of initial (first-mover) migration from villages without recruiters, recruiter location choice should be highly correlated with the extensive margin.[36] To add structure, one can think of recruiters as "traveling salesman" tasked with identifying the least cost method of visiting a set number of locations within a defined area. Suppose that recruitment agencies are required to obtain

[36]Village Potential data from 2008 indicates whether recruiters specifically targeting female migrants visited the given village prior to enumeration. Villages with no migrants in both years have substantially lower recruiter visit rates (1% of villages) than villages with emigrants recorded in both years (23% of villages). See Appendix 1.9.7 for more formal evidence.

operating licenses in district capitals and face a fixed cost of entering villages (e.g., establishing contact with or making royalty payments to village officials). In order to maximize potential migrants reached and minimize fixed entry and variable travel costs, recruiters must first select districts within which to operate and then the order in which villages are visited (see Appendix 1.9.7).

This setup leads to a few testable implications. First, conditional on inter-village travel distance and overall population, districts with fewer villages are more likely to have recruiter visits. Second, the probability of recruiter visits is increasing in district population and decreasing in travel distances between villages within the district. Possible instruments therefore include the district population excluding village v, the number of villages located in v's district, and the area of the district excluding v (as a proxy for inter-village travel distance). Of course, there are reasons that recruiter location choices might directly affect the intensive margin. I address these and related concerns about instrument validity in Section 1.5.5.

1.4.2 Data, Measurement, and Identifying Assumptions

I estimate the equations in (1.10) on a balanced panel comprising agricultural villages in the Village Potential data from 2005 and 2008.[37] I describe here the other data sources used in the empirical analysis and consider a few important identifying assumptions.

Pareto Landholdings Distribution Parameters

Using universal Agricultural Census data from 2003, I follow Gabaix and Ibragimov (2011) to obtain estimates of λ_v for each village by OLS regressions of the log rank(-1/2) on the log landholding size (above $\underline{R} = 0.1$ Ha) using three available measures of assets: (i) total agricultural landholdings, (ii) wetland (or *sawah*) holdings particularly germane to rice production, and (iii) total rice area

[37]The timing of enumeration was fortuitous in that the 2005 round was administered in April preceding the surge in rice prices later that summer. See Appendix 1.9.3 for details on the variables described in this section and Appendix on the panel construction.

planted in 2002-3.[38] The histogram in Figure 1.4 shows the estimated λ_v based on total landholdings to be roughly normally distributed with a slightly fatter right tail of villages with a relatively greater mass of small landholders.

Although the constant minimum bound assumption ($\underline{R}_v \equiv \underline{R}\ \forall v$) serves an important theoretical role along the extensive margin (see Section 1.3.3), it poses empirical challenges because the share of households above $\underline{R} = 0.1$ Ha varies across villages.[39] This is a common problem in the empirical literature comparing size distributions across administrative entities (e.g., Soo, 2005). I pursue a reduced form solution in the empirical analysis in Section 1.5 by controlling for the share of households above $\underline{R} = 0.1$ Ha. Conditional on this measure, which guards against omitted variable bias, $\widehat{\lambda}_v$ captures the most relevant information about the *shape* of the landholdings distribution without neglecting the (near-)landless population. While the Pareto distribution may not provide a good fit in all villages, the two key identifying assumptions are that (i) the share of households below $\underline{R}$ and any departures from Paretian properties above $\underline{R}$ are uncorrelated with the unobservable determinants of migration flows, and (ii) the landholdings distribution in 2003 is predetermined with respect to migration flows between 2005 and 2008 (see Appendix 1.9.6).

Migration Costs

I employ four proxies for observable migration costs: (i) log distance to the nearest city from which labor migrants can feasibly depart Indonesia; (ii) log distance to the (sub)district capital; (iii) the share of Chinese and Arabs in the village as of 2000; and (iv) the share of Muslims in the village. Measures (i) and (ii) capture the most relevant distance-based variation in access to foreign labor markets, while (iii) and (iv) account for differential growth in the demand for immigrant labor across destination countries between 2005 and 2008. Beyond

[38] There is little consensus on the most appropriate method for selecting $\underline{R}$. Clauset et al. (2009) propose a promising approach that nevertheless appears too computationally demanding in the present context. Gabaix (2009) argues that visual inspection should suffice in most cases, and hence I impose $\underline{R} = 0.1$ Ha as the baseline and consider alternatives in robustness checks.

[39] Looking within villages (including semi-urban), the average share of landholding (all) households above 0.1 Ha is 86 (60) percent.

the obvious Arab/Muslim connection to the Middle East, ethnic Chinese may have connections with Hong Kong, Taiwan, and (Chinese in) Malaysia/Singapore. These time-invariant measures will only matter if they capture trends remaining after taking first differences in stock migration rates (see equation (1.10)).

Rainfall Shocks

I employ high resolution rainfall data from the widely used NOAA/GPCP data. Monthly data is mapped into province-specific growing seasons (rather than calendar years) based on the classification provided by Maccini and Yang (2009). The rainfall level in year t corresponds to the total level of rainfall (in centimeters) during a growing season beginning in a latter month of year $t-1$ and ending in mid-t. Two additional stylized facts inform the empirical specification. First, rainfall shocks—measured in terms of log deviations from long-run district-level means—are positively correlated with rice yield shocks (Levine and Yang, 2006; Naylor et al., 2001). Second, there is little evidence that abnormally large rainfall shocks alter the corresponding output gains. In the baseline results, I therefore specify the rainfall shock for 2005 (2008) as the sum of log deviations from long-run district-level means (1953-2008) over the seasons ending in 2003/4/5 (2006/7/8). Figure 1.6 demonstrates the large spatial variation in these shocks across the Indonesian archipelago.

Rice Price Shock

To capture price shocks in the model, I focus on rice, Indonesia's most important agricultural product. I exploit large spatial and time series variation in the domestic price induced by a ban on rice imports beginning in 2004. Initially a temporary policy ahead of the March harvest, the ban was renewed over the next several years in response to sustained political pressure. Prior to 2004, a 20 percent ad valorem tariff had been the primary measure of protection. The ban effectively raised ad valorem rates to around 150 percent, thereby shutting down private sector imports. While rarely exceeding 5 percent of total rice consumption in the decades prior to the ban, imports historically stabilized domestic prices

(Dawe, 2008).

Although the import ban applied universally, the intensity of the subsequent price shock varied considerably across regions. Figure 1.7 demonstrates this by comparing a rice price index across cities throughout Indonesia from January 2002 through March 2008. In Appendix 1.9.5, I motivate and test a simple trade-based model for price changes in local markets that explains some of the vast spatial heterogeneity. The model predicts larger price increases in villages where domestic producers faced greater import competition before the ban. Given prevailing transportation and trade costs, the local import penetration ratio should be decreasing in (i) the distance to the nearest international port and major wholesale markets, and (ii) the shipping distance to the overseas markets from which Indonesia's imports originate. The empirics corroborate these predictions: after the import ban, rice prices grew faster in Indonesian cities more closely aligned with the main rice export shipping routes originating in Thailand and Vietnam.

There are (at least) two explanations for the obvious lack of arbitrage by domestic traders. First, in the wake of decentralization in post-Suharto Indonesia, the state logistics agency (*Bulog*) played a much more limited role in procuring, moving, and equilibrating rice supplies across the archipelago. Second, during the liberal import regime from 1999-2003, private traders developed strong ties with foreign suppliers. The decline of *Bulog* and the path dependence of these private international buyer-seller networks ultimately slowed the process of adjustment to the lack of imported rice. Also, if the import ban led to greater speculative activity in certain domestic markets, then otherwise transitory spikes in local prices might have had longer-lasting effects on future prices than in the absence of speculation (see Deaton and Laroque, 1996).

Before turning to empirical results, I mention two important corollaries to using spatial variation in the price shock to identify the relationship between income changes and migration flows. First, the price shock was not a random discontinuous jump but rather a structural break. If trends broke faster and more sharply in rural areas located closer to central port cities, then any estimated effect of price shocks on migration flows could be biased upward. However, since

I take the (log) difference in migration rates and control for the log distance to the nearest emigration hub, the bias only arises if other *unobservable* migration costs declined more rapidly in villages near port cities. A second concern is that other meaningful economic shocks are incidentally correlated with the rice price shock. The largest of such shocks was presumably an unconditional cash transfer (UCT) equivalent to 120 USD targeted by the government to poor and near-poor households after reducing fuel subsidies in 2005. Since the program effectively reached every village in the country (Bazzi et al., 2012), there is little reason to expect unobserved variation in the local incidence of UCT benefits to be correlated with the rice price shock in such a way as to bias key parameter estimates.

1.5 Empirical Results

This section presents the main empirical results in the paper. I begin by presenting reduced form evidence and then turn to results from the two-step model, which allows us to distinguish between the extensive and intensive margin as suggested by the theory.

1.5.1 Reduced Form

First, I consider estimates of the following specification:

$$\begin{aligned}\frac{M_{v,t+1}}{N_{v,t+1}} &= \theta_p price\ shock_{vt} \ + \ \theta_{p\lambda}(price\ shock_{vt} \times \widehat{\lambda}_v) \ + \ \theta_a rainfall\ shock_{vt} \\ &\quad + \theta_{a\lambda}(rainfall\ shock_{vt} \times \widehat{\lambda}_v) \ + \ \xi_t \ + \ \xi_v \ + \ \varepsilon_{v,t+1}, \qquad (1.12)\end{aligned}$$

where $\widehat{\lambda}_v$ is the estimated shape parameter for total agricultural landholdings; ξ_v (ξ_t) are village (period) fixed effects; the price shock for village v in period t ($t-1$) is the annualized log growth in the abovementioned price index from 2005m4-2008m3 (2002m1-2005m3) in the nearest city;[40] and the sample includes villages with migrants in both 2005 and 2008 as well as those with no migrants in

[40] Relative to unobservable producer prices in local rural markets, the price index should be (i) less affected by supply shocks in small groups of villages, and, (ii) more likely to capture regional general equilibrium effects of the import ban (see Appendix 1.9.5).

one or both years.

Table 1.2 presents estimates of equation (1.12). In columns 1-4, I employ an OLS fixed effects estimator, which is equivalent to a first difference specification. The positive estimates of θ_a and θ_p in column 1 are consistent with liquidity constraints posing a barrier to international migration. At the mean, migration rates increase by 0.18 (1.4) percent for every 1 percent increase in rainfall (price) shocks, though inference is sensitive to the level of clustering (at the village vs. district level). In columns 2-4, price shocks have larger positive effects in villages with a greater mass of small landholders ($\widehat{\theta}_{p\lambda} > 0$), but I cannot reject that the effect of rainfall shocks is constant across $\widehat{\lambda}_v$. Columns 5-8 take a more flexible approach based on the semiparametric Tobit or trimmed least absolute deviations (LAD) (Honoré, 1992), which, unlike OLS, explicitly accounts for the mixture distribution implicit in the cross-section of village migration rates as in equation (1.9). The main qualitative results do not differ substantively from the OLS specification.[41]

Although the positive estimates of the θ parameters in Table 1.2 offer some evidence of liquidity constraints, neither the OLS nor the LAD approach permits inference on whether and to what extent such constraints bind on the intensive and/or extensive margin. These estimators conflate the effects of observable covariates along both margins. Moreover, unobservable differences may also matter, and in order to test the theory, these differences must be made explicit. In the remaining sections, I attempt to overcome such limitations of the reduced form by implementing the two-step approach developed in Section 1.4.

1.5.2 Liquidity Constraints: Baseline Evidence from a Two-Step Model

In this section, I report estimates of the extensive and intensive margins using the two-step model in (1.10). In the second-step, I include 11 fixed effects

[41] Another reduced form approach would be conditional fixed effects Poisson regression with $M_{v,t+1}$ as the dependent variable and $N_{v,t+1}$ as an exposure variable on the right hand side of equation (1.12) with the coefficient on its logarithm constrained to unity. Results using this estimator are qualitatively identical to those in Table 1.2 and are available upon request.

identifying the village's plurality destination country for migrants in 2005.[42] This provides a flexible control for unobservable destination demand shocks that may be common across villages sending migrants to that country. I cluster standard errors at the district level, and all reported second-step significance levels are based on a bootstrap$-t$ procedure (see Appendix 1.9.2).

Extensive Margin—First-Stage

Before proceeding to the intensive margin and tests of Propositions 1 and 2, Tables 1.3 and 1.4 demonstrate three important results for the extensive margin: (i) the role of the landholdings distribution in operationalizing the liquidity and incentive threshold formulation, (ii) the predictive power of the instruments, and (iii) the null effects of rainfall and price shocks. In each table, columns 1-3 report estimates for the extensive margins in 2005 and 2008 based on three discrete choice estimators used in the parametric and semiparametric correction procedures. The parametric correction procedure (Poirier, 1980) requires a bivariate probit first stage. For the semiparametric procedures (Das, Newey and Vella, 2003, hereafter, DNV), I consider a flexible seemingly unrelated linear probability (SU-LPM) (Zellner and Lee, 1965) and semi-nonparametric maximum likelihood (SNP-ML) (Gallant and Nychka, 1987).[43]

In Table 1.3, the robust positive estimate on log maximum landholdings ($\widetilde{R}_v$) and negative estimate on log minimum landholdings ($\underset{\sim}{R}_v$) provide support for the threshold formulation of the extensive margin in equations (1.11). Villages with higher maximum landholdings and lower minimum landholdings (above 0.1 Ha) are more likely, respectively, to have any individuals (i) able to finance migration costs and (ii) with profitable expected income gains from migration. The positive coefficient on maximum landholdings implies that financial constraints bind on the

[42]The 12 plurality destinations (and percent of villages) include: Malaysia (64.9%), Hong Kong (2.8%), Singapore (2.5%), Taiwan (1.6%), Japan (1.3%), South Korea (0.9%), UAE (0.2%), Saudi Arabia (23.5%), Jordan (0.05%), Kuwait (0.3%), USA (1.1%) and Other (0.9%).

[43]Newey (1988) argues that the linear probability model provides consistent estimates in two-step selection models, though a semiparametric first stage estimator provides more efficient (second-stage) estimates (Newey, 2009). Also, as discussed in Appendix 1.9.2, the standard errors in the SNP-ML columns are severely underestimated and should be discounted with respect to the other estimates.

extensive margin. However, there is a downside to including $\widetilde{R}_v$ and $\underset{\sim}{R}_v$ in the first stage. The results in Section 1.3.3 suggest that λ_v and village population size N_v fully determine the expected locations of $\widetilde{R}_v$ and $\underset{\sim}{R}_v$ conditional on the income thresholds R_L and R_U. Given the central role of λ_v in the second stage, I therefore retain the more general extensive margin specification with N_v and $\widehat{\lambda}_v$ in the first stage moving forward.

Table 1.4 reveals robust negative point estimates on $\widehat{\lambda}_v$. This suggests that the probability that village v crosses the extensive margin thresholds is increasing in the share of large landholders within v. Evaluating the bivariate probit estimates at the mean in 2005, for example, a 25 percent reduction in $\widehat{\lambda}_v$ (to its 25th percentile) increases the probability of having any emigrants from 0.59 to 0.66. These effects are again consistent with liquidity constraints mattering more than incentive constraints on the extensive margin: in the typical zero migration village v, all households fall below the minimum wealth requirement $\widetilde{R}_v$ (liquidity threshold) rather than above the minimum expected income differential $\underset{\sim}{R}_v$ (incentive threshold).

Table 1.4 offers a few additional results of interest. First, rainfall and rice price shocks have positive albeit statistically insignificant effects on the extensive margin. This may explain some of the muted reduced form results in Table 1.2. Second, geographically remote villages are less likely to have any emigrants. Lastly, the point estimates for log village v population and log district population and area less v conform with the traveling salesman framework discussed in Section 1.4.1. Insomuch as these instruments are isolating excludable variation in recruiter presence, the local average treatment effect of income shocks can be used to identify the presence of liquidity constraints in those villages induced into international labor markets by the presence of recruiters. From a migration policy perspective, this is precisely the treatment effect of interest.

Intensive Margin—Second-Stage

Table 1.5 presents estimates of the following second-stage specification aimed at testing Proposition 1:[44]

$$\Delta \ln \left(\frac{M_{v,t+1}}{N_{v,t+1}} \right) = \theta_a \Delta rainfall\ shock_{vt} + \theta_p \Delta price\ shock_{vt} \\ + \chi \widehat{\lambda}_v + \boldsymbol{\zeta}' \mathbf{X}_{vt} + f(\widehat{P}_{v,t+1}, \widehat{P}_{vt}) + \Delta \varepsilon_{v,t+1} \ , \quad (1.13)$$

where $\widehat{P}_{v,t+1}$ and $\widehat{P}_{vt}$ are estimated correction terms based on first stage specifications with the instruments as in Table 1.4. I report estimated parameters using $\widehat{\lambda}_v$ based on total agricultural landholdings, wetland holdings, and rice paddy area planted in 2002-3. In the parametric Poirier procedure, the $f(\widehat{P}_{v,t+1}, \widehat{P}_{vt})$ function is the sum of two (bivariate) Mills ratios. In the Das et al. procedure, $f(\cdot)$ is a 3rd order polynomial in the propensity scores, but the results are robust to other orders or functional forms (e.g., 25 bins in the propensity scores). In all cases, the correction terms (suppressed for presentational purposes) are jointly statistically significant and are not highly correlated with other second-stage covariates.

The first main result is that rainfall shocks have a statistically significant positive effect on flow migration rates. Focusing on the correction-adjusted estimates in columns 2-4 of the top panel (total agricultural landholdings), we find that a 10 percent increase in cumulative rainfall shocks between periods implies roughly a 3-4 percent increase in flow migration rates. These are economically meaningful effects: at the mean, a one standard deviation increase in inter-period rainfall shocks moves the log flow migration rate from 0.11 to the 75th percentile of 0.69. Restricting to $\widehat{\lambda}_v$ based on wetland holdings (columns 6-8) and paddy area planted (columns 10-12) in the bottom two panels, we also find positive albeit slightly smaller elasticities.

Rice price shocks also have a positive effect on flow migration rates, but the estimates $\widehat{\theta}_p$ vary in magnitude and significance. That the price shock retains

[44]The time-invariant covariates (including $\widehat{\lambda}_v$) are strictly not necessary in the second stage but are included so as to allow the effect of such variables to differ across periods. The estimates for these other variables have mostly been suppressed from the tables but are available upon request. All results are robust to using a more parsimonious specification (see Section 1.5.5).

significance conditional on the rainfall shock confirms that prices contain information beyond rainfall volatility, including the effects of the import ban. The upper bound estimate implies that a 1 percent increase in the annualized price shock between periods is associated with a 3.5 percent increase in the flow migration rate. The elasticities are relatively smaller when using the bivariate probit or SU-LPM first-stage estimators. By allowing for correlation across periods in the extensive margin equations, the correction terms in these columns remove more of the independent variation in the effect of the price shock on the intensive margin. I retain these more flexible (and conservative) first-stage estimators moving forward.

In the context of Proposition 1, the estimates of θ_a and θ_p in Table 1.5 suggest that financial constraints to international migration are binding on average across villages in rural Indonesia. In other words, the lower liquidity threshold R_L in equation (1.2) is binding in the typical village (i.e., there is nonzero population below R_L). The positive and mostly statistically precise coefficients imply that the average village is in the liquidity-constrained regime where the upfront costs of migration prevent the uptake of profitable international migration opportunities for some individuals.

However, the smaller and imprecise estimates of θ_p and θ_a in columns 1, 5 and 9 of Table 1.5 suggest that ignoring the extensive margin understates the importance of these financial constraints on the intensive margin. By definition, the villages in Table 1.5 have some households capable of crossing the two extensive margin thresholds. Empirically, Tables 1.3 and 1.4 showed, using landholdings distribution statistics, that financial constraints are more binding than profitability constraints along the extensive margin. Thus, the correction terms account for the fact that villages in Table 1.5 have relatively fewer liquidity-constrained households than zero migration villages not in the Table. This makes it possible then to use exogenous agricultural income shocks to test whether financial constraints bind on the intensive margin, inhibiting migration among the poor in villages with established connections to international labor markets.

1.5.3 Landholdings Heterogeneity and the Effects of Rainfall and Price Shocks

Having found robust positive elasticities of migration flows with respect to rainfall (θ_a) and rice price shocks (θ_p), I demonstrate in this section how the landholdings distribution affects these parameters. In the previous Table 1.5, we see that lower dispersion in landholdings (high $\widehat{\lambda}_v$) is associated with higher flow migration rates. Yet, we found the opposite along the extensive margin estimates in Table 1.4. Here, I show that this important difference can be explained by agricultural income shocks inducing greater international migration in villages above the extensive margin thresholds that have relatively more households in the lower tail of the landholdings distribution.

I begin by testing the rainfall predictions in Proposition 2 using the following specification:

$$\begin{aligned} \Delta \ln \left(\frac{M_{v,t+1}}{N_{v,t+1}} \right) &= \chi \widehat{\lambda}_v + \theta_a \Delta rain\ shock_{vt} + \theta_{a\lambda}(\Delta rain\ shock_{vt} \times \widehat{\lambda}_v) \\ &\quad + \boldsymbol{\zeta}' \mathbf{X}_{vt} + f(\widehat{P}_{v,t+1}, \widehat{P}_{vt}) + \Delta \varepsilon_{v,t+1}. \end{aligned} \tag{1.14}$$

Assuming the model identifying assumptions hold, the estimates of $\widehat{\theta}_{a\lambda}$ in Table 1.6 should be viewed as lower bounds for the true $\theta_{a\lambda}$ insomuch as the λ_v are estimated with any error (in the auxiliary OLS regressions).

Columns 1-9 in Table 1.6 provide strong evidence that rainfall shocks have larger positive effects in villages with lower mean and less inequality in landholdings. The baseline $\widehat{\theta}_{a\lambda} = 0.147$ in column 2 using total agricultural landholdings implies that villages with $\widehat{\lambda}_v$ one standard deviation above the mean (with landholdings $mean_v = 0.18$ Ha, $Gini_v = 0.30$) have elasticities of flow migration rates with respect to rainfall shocks equivalent to 0.5 whereas villages one standard deviation below the mean (with landholdings $mean_v = 0.28$ Ha, $Gini_v = 0.47$) have elasticities around 0.3. These are small but economically meaningful differences. The estimates of $\theta_{a\lambda}$ slightly differ across parametric and semiparametric corrections but are largely invariant to the type of landholdings used to estimate λ_v. Overall, the evidence again points to liquidity constraints being an important barrier to international migration flows.

In the bottom panel of Table 1.6, I augment specification (1.14) with the interaction of $\widehat{\lambda}_v$ and the price shock:

$$\begin{aligned}\Delta \ln\left(\frac{M_{v,t+1}}{N_{v,t+1}}\right) &= \chi\widehat{\lambda}_v + \theta_a \Delta rain\ shock_{vt} + \theta_{a\lambda}(\Delta rain\ shock_{vt} \times \widehat{\lambda}_v) \\ &\quad + \theta_p \Delta price\ shock_{vt} \ + \ \theta_{p\lambda}(\Delta price\ shock_{vt} \times \widehat{\lambda}_v) \\ &\quad + \boldsymbol{\zeta}'\mathbf{X}_{vt} + f(\widehat{P}_{v,t+1}, \widehat{P}_{vt}) + \Delta\varepsilon_{v,t+1}. \qquad (1.15)\end{aligned}$$

The correction-adjusted estimates of $\theta_{p\lambda}$ in columns 10-18 show that the elasticity of flow migration rates with respect to price shocks is higher in villages with a greater mass of small landholders. However, unlike rainfall shocks, the type of landholdings matters. Though positive, the estimates of $\theta_{p\lambda}$ in columns 2 and 3 using total agricultural landholdings are small and statistically insignificant. Restricting to landholdings specific to rice production in columns 4-9 yields large and statistically significant positive estimates $\widehat{\theta}_{p\lambda}$. That such patterns did not arise for rainfall shocks may be due to the fact that several other crops besides rice are dependent on rainfall. More interestingly, it suggests that the empirical tests are indeed picking up variation in binding financial constraints. Expected incomes are rising fastest for net rice producers, and yet we observe the largest, positive heterogeneous effects of price shocks on migration when removing non-rice farmers (i.e., net consumers for whom rice price increases have a negative effect on real income) from the estimation of λ_v.[45] These effects are also economically meaningful. Taking the value of $\widehat{\theta}_{p\lambda} = 1.155$ in column 6, for example, villages with $\widehat{\lambda}_v$ at the 75th percentile exhibit an elasticity of flow migration rates with respect to price shocks equivalent to 0.9 whereas villages with $\widehat{\lambda}_v$ at the 25th percentile have elasticities around 0.3.

Applying the quasi-structural interpretation of $\theta_{p\lambda} = 1/\beta$, I find in column 14 that $\widehat{\beta} = 0.52$ for wetland holdings.[46] This finding is consistent with available

[45] In unreported results, I also find that rainfall and rice price shocks have larger positive effects in villages with large rice mills, the presence of which are associated with net rice production.

[46] Given plausible attenuation bias in $\widehat{\theta}_{p\lambda}$ arising from measurement error in $\widehat{\lambda}_v$, I take the largest estimate in Table 1.6.

estimates in the agricultural literature. In Bazzi (2012b), I estimate $\widetilde{\beta} = 0.55$ using auxiliary household survey data on wetland holdings and rice output from 2004. Moreover, applying the delta method to $1/\widehat{\theta}_{p\lambda}$, I reject at the 95% level that $\widehat{\beta} = 0.34$, the lowest estimate of β in the literature on Indonesian agriculture (Fuglie, 2010b), but I fail to reject that $\widehat{\beta} = 0.69$, the largest available estimate (Mundlak et al., 2004).

In Table 1.7, I augment the specifications in Table 1.6 by allowing the effect of rainfall and price shocks to vary with the population share of (near-)landless and paid agricultural laborers. In columns 1-9, we find that the elasticity of migration flows with respect to rainfall shocks is declining in the share of (near-)landless and paid agricultural laborers. Controlling for these additional population characteristics leads to lower, albeit still positive, estimates of $\theta_{a\lambda}$. Although the theory does not offer explicit guidance on how to incorporate the (near-)landless, the estimates in columns 10-18 provide complementary insights beyond the Pareto distribution. Paid agricultural laborers typically hail from the poorest landless households in which international migration is least likely (see Figure 1.2). One explanation for the results, then, is that, conditional on the heterogeneity with respect to landholdings dispersion above 0.1 Ha, the positive effect of income shocks on migration flows may be decreasing in the share of (near-)landless, net rice consuming households below 0.1 Ha for whom liquidity constraints are most likely to bind unconditionally.[47]

Turning to columns 10-18, price shocks have (i) larger effects in villages with a greater mass of small landholders above 0.1 Ha, (ii) smaller effects in villages with a higher share of households below 0.1 Ha, and (iii) larger effects in villages with relatively more paid agricultural laborers. The results in (i) and (ii) are consistent with the analogous argument outlined above with respect to rainfall

[47] A related possibility is that the landless population in *semi*-rural areas includes many non-poor households—as borne out in household survey data—for whom transitory rainfall shocks have little effect on income. An alternative explanation is that households with low landholdings above 0.1 Ha can now afford to send a member to work abroad and subsequently must hire labor out of the (near-)landless population. In this case, agricultural wages—already increasing in response to positive rainfall shocks—are likely to rise even further and hence disincentivize migration for (near-)landless households, which for unobservable reasons (unrelated to landholdings) may have been able to afford migration costs *ex ante*.

shocks. The finding in (iii) merits an alternative explanation. It is possible that the wage gains from the price shock, accumulated over multiple growing seasons, generated sufficiently large income gains for (near-)landless paid agricultural laborers to afford the costs of migration. Although outside the explicit theoretical framework in this paper, this explanation would nevertheless be consistent with its intuition.

Taken together, the results in this section suggest an important role for the landholdings distribution in determining the aggregate effect of income shocks on migration flows. In the context of the theory, the statistically and economically significant estimates of $\widehat{\theta}_{p\lambda}$ and $\widehat{\theta}_{a\lambda}$ are consistent with widely prevalent financial barriers to international migration. Moreover, to the extent that λ_v is estimated with error, these estimates may understate the true extent of such financial constraints in the population.

1.5.4 Further Evidence of Liquidity Constraints

This section provides further evidence pointing towards financial constraints along the intensive margin. The empirical tests are based on the idea that, conditional on the distribution of landholdings, transitory income shocks should have smaller effects on flow migration rates in villages with higher prevailing wealth levels. Table 1.8 tests these hypotheses using the following specification:

$$\begin{aligned} \Delta \ln\left(\frac{M_{v,t+1}}{N_{v,t+1}}\right) &= \theta_z z_{vt} + \theta_a \Delta rain\ shock_{vt} + \theta_{az}(\Delta rain\ shock_{vt} \times z_{vt}) \\ &\quad + \boldsymbol{\zeta}'\mathbf{X}_{vt} + f(\widehat{P}_{v,t+1}, \widehat{P}_{vt}) + \Delta\varepsilon_{v,t+1}, \end{aligned} \tag{1.16}$$

where z_{vt} is a proxy for wealth. Under binding financial constraints, we should observe $\theta_{az} < 0\ \forall z$.

I consider four measures of wealth (z). First, I use bank presence at baseline. Given profit maximizing behavior of most banks operating in rural Indonesia, I view their presence in the village or subdistrict as indicative of higher wealth levels.[48] Bank presence also increases the probability that individuals with sufficient

[48] I observe whether Bank Rakyat Indonesia (BRI), rural people's banks (Bank Perdesaan

collateral are able to obtain loans, and recruiters active in the area can secure credit more easily for subsequent on-lending to potential migrants. Second, I utilize a village-level estimate of mean household expenditures per capita in 2000 based on the Elbers et al. (2003) poverty mapping methodology, implemented for Indonesia in Suryahadi et al. (2005), and containing no information about household landholdings. Third, I consider a binary indicator for whether any land in the village was technically irrigated in 2005. Technical irrigation systems often provide sufficient water for rice production even in the absence of requisite rainfall levels during the wet season. Finally, I consider the total rice output per hectare planted in the village in 2001 before the sample period as a measure of aggregate productivity that is informative about prevailing inter-village wealth differentials.

Columns 1-12 of Table 1.8 provide consistent evidence that rainfall shocks have larger positive effects on migration flows in more economically underdeveloped villages. First, bank presence in the village's subdistrict is associated with statistically and economically significantly lower elasticities of migration flows with respect to rainfall shocks. In column 2, for example, elasticities fall from 0.57 to 0.35 in villages with access to financial institutions. Second, $\widehat{\theta}_{az} \approx -0.875$ in columns 5-6, which implies that villages at the 25th percentile of log mean household expenditures per capita have an elasticity of migration flows with respect to rainfall shocks equivalent to 0.95 compared to an elasticity of 0.66 in villages at the 75th percentile. Third, in columns 8-9, the elasticity doubles in moving from villages with technical irrigation systems to those without. Fourth, in columns 11-12, villages at the 25th percentile of log aggregate rice productivity have an elasticity of 0.34 relative to an elasticity of 0.25 for villages at the 75th percentile.

1.5.5 Robustness Checks and Alternative Explanations

The previous sections highlighted several empirical results that are broadly consistent with the widespread prevalence of binding financial constraints to inter-

---

Rakyat or BPR), or formal commercial banks operate in the village. Although BRI location decisions may be orthogonal to pre-existing wealth levels (see Gertler et al., 2009), I retain the assumption that a broader measure of bank locations is informative about prevailing wealth levels.

national migration from Indonesia. In Appendices 1.9.4, I take additional steps to show that these key results are generally robust to and in some cases strengthened by the following conditions:

(D.1) controlling for the effects of natural disasters

(D.2) controlling for other agricultural commodity price shocks or overall regional agricultural GDP shocks

(D.3) alternative specifications for and measurement of the rainfall and price shocks

(D.4) alternative choices of $\underline{R}$ in the estimation of λ_v

(D.5) imposing a more parsimonious vector of time-invariant, second-stage covariates

(D.6) controlling for past internal migration, demographic structure, and average household size

(D.7) accounting for outliers (in M_{vt} and $\widehat{\lambda}_v$) and the quality of village population registers

This battery of robustness checks increases our confidence in interpreting the main empirical results through the lens of the theory. In the remainder of this section, I aim to rule out other alternative explanations and reconsider instrument validity.

Ruling Out Other Alternative Explanations

I consider here three potential confounding factors in the empirical analysis and rule out their having substantive effects on the results. One concern could be that demand shocks or policy barriers in destination countries have differential effects on emigration rates across villages. There are two reasons that these shocks should not matter for the results. First, prevailing visa policies do not discriminate by region of origin in the home country, and by controlling for the ethno-religious composition of villages, I account for other discriminatory policy barriers (e.g., if Saudi Arabia screens on religion). Second, I control for each village's baseline plurality destination, which should capture differential sorting and exposure to demand shocks.

A second concern is that unexplained sources of idiosyncratic heterogeneity in migration costs or preferences systematically vary across villages. By taking first differences, I remove any time-invariant differences in these unobservables across villages. After conditioning on ethno-religious composition, there is little reason to suspect large between-village differences in the within-village distribution of preferences. Moreover, given the small size of Indonesian villages, potential network externalities in migration costs (see Carrington et al., 1996) are more likely to be fixed than to vary across individuals within a given village. Also, in all specifications, I account for skill heterogeneity by controlling for the share of individuals in the village with post-primary education in 2000. In results that are available upon request, the effect of agricultural income shocks does not vary with the schooling distribution.

A third concern is that although agricultural income shocks may affect internal emigration flows, which are unobserved at the village level, the plausible effects work against the key empirical results. Recall that the log flow migration rate for village v is defined as $\widehat{M}_{v,t+1} \equiv \Delta \ln \left(M_{v,t+1}/N_{v,t+1}\right)$, and international emigrants are defined as village residents and hence included in $N_{vs} \forall s$. Insomuch as unobservable internal migrants affect population records (N_v), differences in $\widehat{M}_{v,t+1}$ across villages could be driven by internal rather than international migration flows. This could bias our interpretation of the main empirical results. However, two factors suggest that the direction of bias goes against my findings. First, I find that positive rainfall shocks at home reduce internal out-migration using available district-level panel data constructed from population censuses (see Appendix 1.9.4).[49] Second, using multiple data sources, Hugo (2000) argues that favorably endowed rural areas receive immigrants during periods of high agricultural commodity prices. Both mechanisms imply that positive agricultural income

[49]Formally, rainfall shocks ($shock^a$) lead to a proportional change in the number of international migrants that is larger than the proportional change in village population: $\partial \Delta \ln M_{v,t+1} / \partial \ln shock^a_{vt} > \partial \Delta \ln N_{v,t+1} / \partial \ln shock^a_{vt}$. Kleemans and Magruder (2012) also find that positive rainfall shocks reduce rural-to-urban migration among individuals in the Indonesian Family Life Survey dataset.

shocks increase village population size. Thus, for a given change in the number of international migrants, unobservable internal migration flows attenuate the estimates of the true parameters estimated in Tables 1.5-1.7.

A Further Look at Instrument Validity

Though there is little concern with weak instruments (see Tables 1.3 and 1.4), the validity of the exclusion restrictions in the two-step model requires further discussion. As in all Heckman-type models, it is not straightforward to derive closed-form expressions for the possible biases arising from invalid instruments, or to apply new tests for instrument validity designed for linear 2SLS models (e.g., Conley et al., forthcoming). Nevertheless, the primary results do not rest entirely on the associated excludability assumptions.

Table 1.9 varies the exclusion restrictions employed in estimating the key parameters $\widetilde{\Theta} \equiv (\theta_a, \theta_p, \theta_{a\lambda}, \theta_{p\lambda})$ using the Das et al. (2003) procedure and $\widehat{\lambda}_v$ based on wetland holdings. With four instruments and two first stage equations, I can assess the effect of treating at most two instruments as non-excludable. In the table, I compare baseline estimates of $\widetilde{\Theta}$ in columns 1-3 to estimates when including (i) the log number of villages in v's district (columns 4-6), (ii) the log number of villages in v's district and the log area of v's district less v (columns 7-9), (iii) the log population of v (columns 10-12), (iv) the log population of v and area in v's district less v (columns 13-15), and (v) the log population of v and v's district less v (columns 16-18). Aside from a few insignificant differences, I find no systematic departures from the baseline results.

Though encouraging, the results in Table 1.9 can only provide a partial test of instrument validity. Nevertheless, some potentially salient correlations between instruments and unobservables in the second stage $\Delta\varepsilon_{v,t+1}$ should work against my main findings. By expressing the second stage in log differences and full elasticity form, I remove some of the time-invariant level differences in labor market size and density (and by proxy, wage differentials) across regions. This addresses the concern that the instruments are merely identifying variation in the vitality of local labor markets relevant to village v residents. Another relevant concern is that the instruments are positively correlated with unobservable village-level migration networks. Yet, this would imply smaller effects of income shocks because for villages induced into the second stage, the informational costs of migration are relatively lower, and (the relatively wealthier) past migrants can also provide informal financial support.

1.5.6 From Migration Choices to Migration Flows: A Validation Exercise[50]

In this section, I use auxiliary micro data to validate the mapping from liquidity-constrained migration choices to aggregate migration flows implied by the village-level regressions. I proceed in two steps. First, I estimate a migration choice model using nationally representative household survey data (*Susenas*) collected in mid-2006.[51] I then relate the heterogeneous marginal effects of income shocks in these regressions to the village-level elasticities reported in Section 1.5.3.

In columns 1-2 of Table 1.10, I report average marginal effects (AMEs) of agricultural income shocks on migration choice between 2000 and 2006 based on variants of the following specification:

$$\begin{aligned} \mathbb{P}(migrate_{iv,t+1} = 1) &= F(\beta \; rainfall \; shock_{vt} + \; \gamma \; price \; shock_{vt} \\ &+ \quad \eta_i + \; \eta_t + \; e_{iv,t+1} > 0) \; , \end{aligned} \tag{1.17}$$

using a conditional fixed effects logit estimator (CFE-logit) where $migrate_{iv,t+1} = 1$ if household i in village v had any migrants depart in year $t+1$, η_i (η_t) are household (year) fixed effects, and $e_{iv,t+1}$ is an idiosyncratic error term. Note that the CFE-logit estimator restricts estimation to ever-migrant households (i.e., with $migrate_{iv,t+1} = 1$ for at least one t). The positive and precisely estimated AMEs suggest that positive agricultural income shocks increase the probability that households send members to work abroad next period. In columns 3-4, I

[50] Further details on the analyses in this section can be found in Appendix 1.9.4.

[51] The *Susenas* data are collected from around 10,000 households in 670 villages and elicit information on the occupation and date of departure of household members that ever worked or are currently working abroad.

allow the effect of income shocks to vary with household wetland holdings and find that both rainfall and rice price shocks have the largest positive effects among small landholders. In the more flexible quadratic specification, the implied AMEs, which I estimate at each landholding size in the sample, are positive and precisely estimated for households with landholdings less than 0.6 Ha but are negative and/or insignificant for larger landholders. Lastly, column 5 shows that rainfall shocks have larger positive effects on household migration choice in villages without nearby banks.

In light of the microfoundations of the theoretical model, the results in Table 1.10 are indicative of binding liquidity constraints in migration choice. Moreover, the findings are qualitatively consistent with the main village-level results showing that agricultural income shocks have larger positive effects on international migration flows in villages with a greater mass of small landholders. This similarity is reassuring in that the household-level regressions provide evidence against aggregation bias. However, we can use the estimates in Table 1.10 to go a step further in evaluating the quality of the mapping from individual choices to aggregate outcomes suggested by the theory and operationalized empirically in the village-level regressions.

Using the estimated AMEs from columns 3 or 4 and the village-specific Pareto parameters $\widehat{\lambda}_v$, I construct village-level elasticities of migration flows with respect to rainfall and price shocks. For each village v, I first assign the nationally representative AMEs (for each shock) to the share of households at the given landholding size implied by $\widehat{\lambda}_v$. I then sum those weighted AMEs across the village population to recover an elasticity of village migration *flows* with respect to income shocks that can be compared to the actual village-level elasticities obtained from $\widetilde{\Theta}$ parameters in previous regressions based on equation (1.15).

Rendering the estimates comparable across the micro and macro data in this way, I show in Table 1.11 that the elasticities based on the reweighted AMEs from Table 1.10 (either linear or quadratic) are not only highly correlated with but also have similar means and medians as the elasticities from the village-level regression in column 14 of Table 1.6. These similarities corroborate the prevalence of financially-constrained migration choices that were inferred from the main village-level regressions using the theory. However, Table 1.11 also shows that the village-level regressions yield elasticities with much greater variance across villages than those based on aggregating the individual AMEs. This key difference is informative about the importance of fixed costs in the context of the model. In the aggregation of the individual-level AMEs, the effect of income shocks at landholding size $R_{iv} = R'$ is identical across villages (by construction). On the other hand, in the village-level regressions, the underlying (unobservable) effect of income shocks on migration choice for individuals with $R_{iv} = R'$ may differ across villages depending on village-specific migration costs. Large variation in these costs across Indonesia could explain the relatively higher variance of the elasticities from the village-level regressions. I proceed next to document precisely such variation.

1.5.7 Village-Specific Migration Costs

In this final section, I document systematic variation in migration costs across the Indonesian archipelago. Having found robust empirical evidence pointing towards financial constraints, I use equation (1.4) to recover lower bounds for village-specific costs associated with prevailing two-year labor contracts. The two-step regression framework allows me to derive these costs for all villages including those with no migrants.

I begin by describing the analytical procedure. First, using all second-step parameters from the baseline specification in column 14 of Table 1.6, I predict the log flow migration rate, $\Delta \ln (M_{v,t+1}/N_{v,t+1})$, for all villages. Second, I recover the Cobb Douglas coefficient on landholdings $\widehat{\beta} = 0.52$ (see Section 1.5.3). Third, using the regional monthly rice price indices, I estimate village-specific autoregressive parameters α_v at a bi-annual frequency.[52] Third, I plug in the appropriate empirical analogues for rice price shocks $\Delta \ln p_{vt}$ and rainfall $\overline{\sigma}_v + a_{vt} = \sigma_{vt}$. Fourth, I use the destination-specific monthly gross wages reported by Bank Indonesia (2009) to calculate the wage offers ω_{vjt} most relevant to village v residents in period t. For

[52]The bi-annual frequency is a plausible time frame over which households (i) forecast prices into the next of the two growing seasons over the agricultural year, and (ii) make decisions over temporary migration opportunities.

villages with any migrants in 2005, ω_{vjt} equals the two-year gross wage offered to Indonesians around 2005 in the plurality destination of migrants from that village. The wage in period $t+1$ equals the two-year gross wage in 2008 in that same destination.[53] For villages with no migrants in 2005, ω_{vjt} equals the average value among villages with any migrants in their district. I then solve analytically for migration costs C_{vjt} in equation (1.4) after a few simplifications (see Appendix 1.9.4).

Table 1.12 reports summary statistics for the estimated migration costs in roughly 42,000 Indonesian villages. These costs range from 100 to 5,200 USD and are around 1,500 USD for the average village.[54]

There are at least two reasons why the migration costs implied by the structural model can exceed the destination-specific recruitment and placement fees quoted by the Indonesian government (ranging from 800-1,200 USD in Bank Indonesia, 2009). First, the estimated costs include potential interest that accrues on any pre-departure loans advanced to migrants either by recruitment agencies or other (informal) lenders. World Bank (2009) provides evidence that effective annual interest rates on loans from recruiters range from 50-60 percent. Given that the typical migrant only pays around 40-50 percent of total fees upfront, the resulting effective costs can certainly surpass 1,500 USD for some destinations.

Second, and more importantly, the estimated costs include potential search or information costs as well as differences in the travel costs of reaching international departure points across the archipelago. This can be seen in Figure 8(a), which plots average village-specific costs by district. Costs appear to be relatively lower in (i) areas of Java and South Sulawesi that are better connected to international air transport hubs, (ii) the eastern coast of Sumatra and the western coast

[53] Bank Indonesia (2009) lists gross nominal wages by destination as agreed upon in bilateral Memoranda of Understanding and posted by recruiters. Reported wages in their survey exhibit little variation around these officially set wages. Average monthly wages range from 130-180 USD in Malaysia/Singapore to 175 USD in the Middle East to 400-550 USD in Hong Kong and Taiwan.

[54] The variance in costs is relatively low on account of imposing identical gross wages for all individuals going to a given destination. Generalizing to a case where villages and households draw randomly from a lognormal gross wage distribution leaves the qualitative patterns unchanged without generating additional insights.

of Kalimantan, which are a short distance over water to peninsular Malaysia and Singapore, and (iii) areas of East Nusa Tenggara with a long history of (undocumented) migration to Malaysia (Hugo, 2008).

Although the estimated costs constitute a relatively large financial burden on the typical Indonesian household, the implied potential net income gains to migration are also quite substantial. For the average household, migration costs are equivalent to 1.5 years of total expenditures (as a proxy for permanent income).[55] The map in Figure 8(b) shows considerable variation in this ratio across Indonesia. Costs are relatively lower as a share of expenditures in areas of Kalimantan and Sumatra near Malaysia/Singapore, and rural areas near major urban centers with high earnings potential in Java and South Sulawesi. For the typical Indonesian household, however, the implied net earnings from working abroad for two years ($\omega_{vjt} - C_{vjt}$) are nearly 35 percent larger than cumulative household expenditures over the same period. Figure 8(c) shows that this differential ranges from 50-170 percent in most areas of densely-populated rural Java and in East Nusa Tenggara, among the poorest regions of Eastern Indonesia. Overall, these estimates suggest large interregional differences in potential future emigration flows.

1.6 Conclusion

In this paper, I develop a novel theoretical framework that makes it possible to identify the effect of changes in individual incomes on international migration flows. Drawing upon a rich empirical context in Indonesia, I provide new evidence on the extent to which financial constraints shape international migration flows in low-income settings. Consistent with theoretical predictions identifying such constraints, positive rainfall and rice price shocks are associated with greater international migration, particularly in villages with a greater mass of small landholders. Using auxiliary household-level data, I validate the microfoundations of

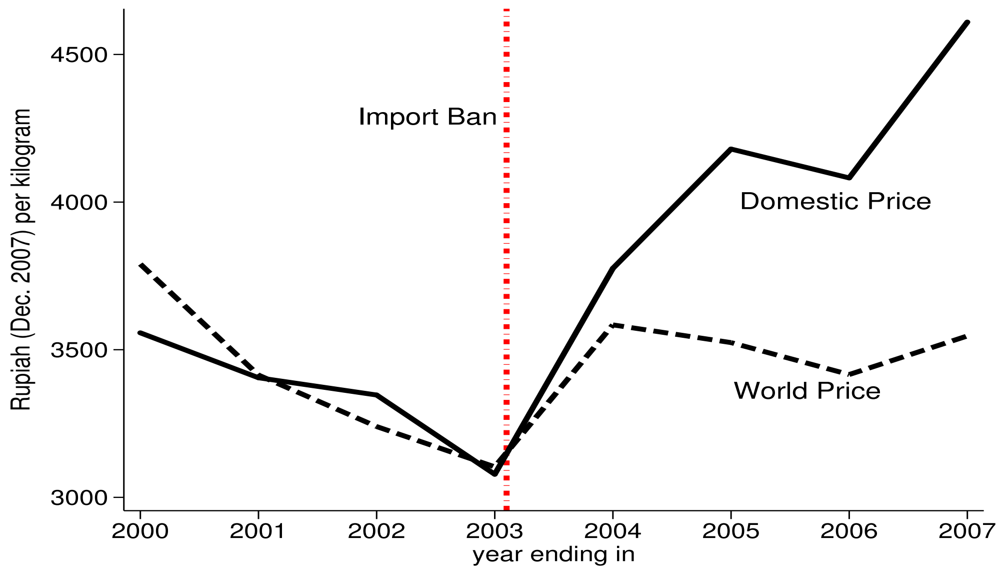
[55] This estimate is based on expenditure data from household surveys representative at the district level from 2006-7. As in other developing countries, expenditure data is more reliable than earnings data and is typically a better measure of permanent income. Similar results are obtained when using district-level GDP per capita.

liquidity-constrained migration choices underlying the main village-level results. Lastly, I estimate lower bounds on village-specific migration costs and document considerable variation across Indonesia in the magnitude of these costs as well as in potential income gains to working abroad. Overall, the results offer a new window into the ongoing debate over the implications of rising incomes in developing countries for global migration flows.

Whether the empirical results in this paper extend to other migration channels and developing countries is an important question for future research and one to which the theory in this paper can be readily adapted. The theory offers a new way to formulate gravity-based specifications of migration flows in the presence of zeros, which are common in popular migration datasets.[56] One interesting extension in this regard would be to allow for multiple destinations (including internal) with different fixed costs. Moreover, beyond the agricultural context, one could extend the model to a longer time horizon and other dimensions of income heterogeneity, which may play an important role in governing the effects of biased technical change on migration flows. More generally, the aggregation procedure put forward in this paper could be applied to other economic problems in which wealth heterogeneity matters and for which existing mappings from individual choices to macro outcomes may be empirically intractable.

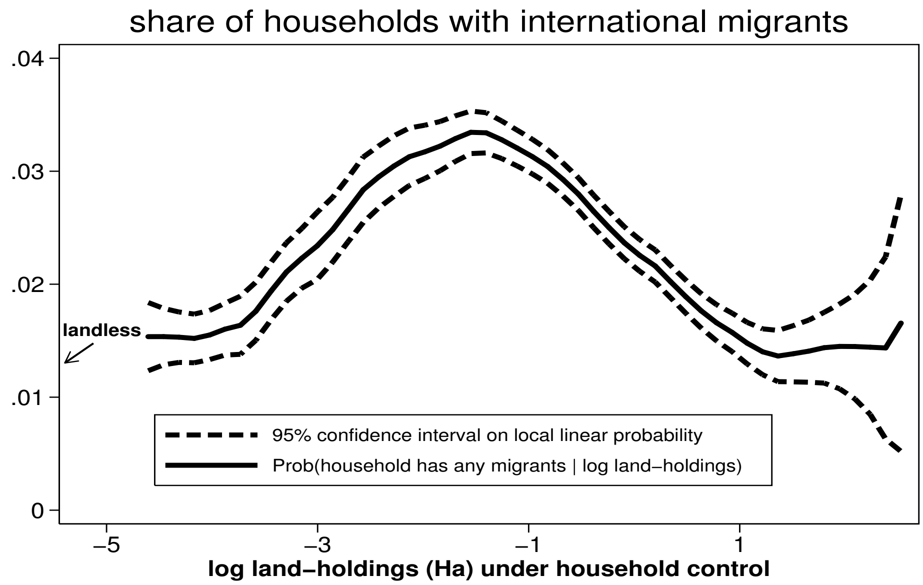
[56] In the widely used Global Migrant Origin Database, for example, over 25 percent of 226×225 bilateral pairs of countries have no foreign-born individuals based on the 2000 round of Population Censuses; Beine et al. (2011) report zeros in approximately 35 percent of bilateral pairs in their study of diasporas using data from 1990 and 2000; roughly 95 percent of origin country $\times$ destination US state observations are zero in Simpson and Sparber (2010).

1.7 Figures



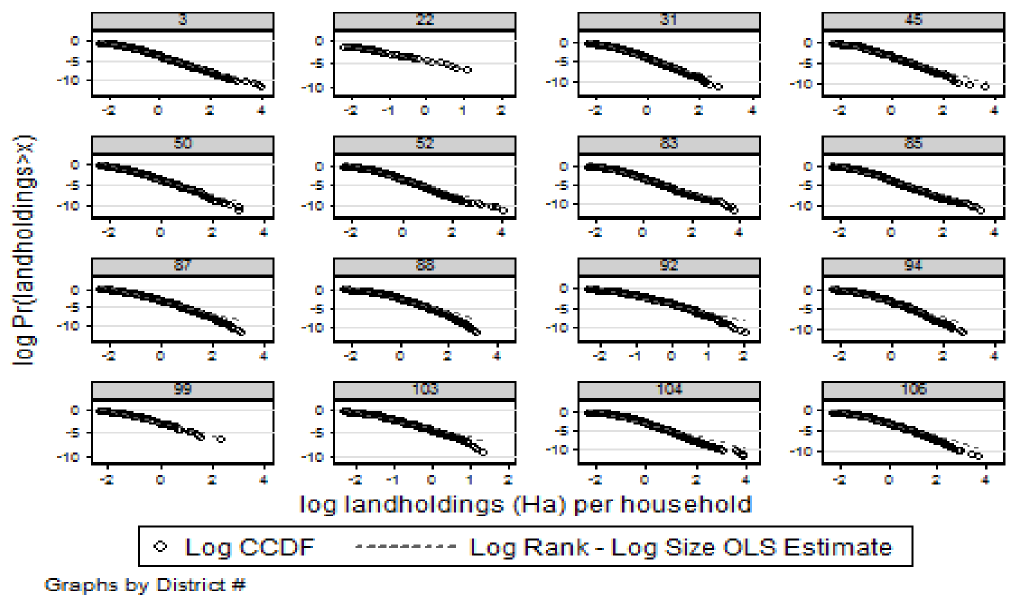
Notes: Year-end average farmgate/producer prices from 2000 to 2007 across Indonesia reported by the Food and Agriculture Organization (FAO). Nominal prices are deflated by the national CPI reported by Bank of Indonesia. Exchange rate and world price data are obtained from the IMF. Further adjustments are made as suggested in Dawe (2008): Thai 100B f.o.b. adjusted to retail level by USD 20 per ton and 10% markup from wholesale to retail, adjusted downward for quality by 20% from 1991-2000 and by 10% from 2001-2007 based on trends in quality preferences in the world market.

Figure 1.1: World vs. Domestic Rice Prices (year-end)



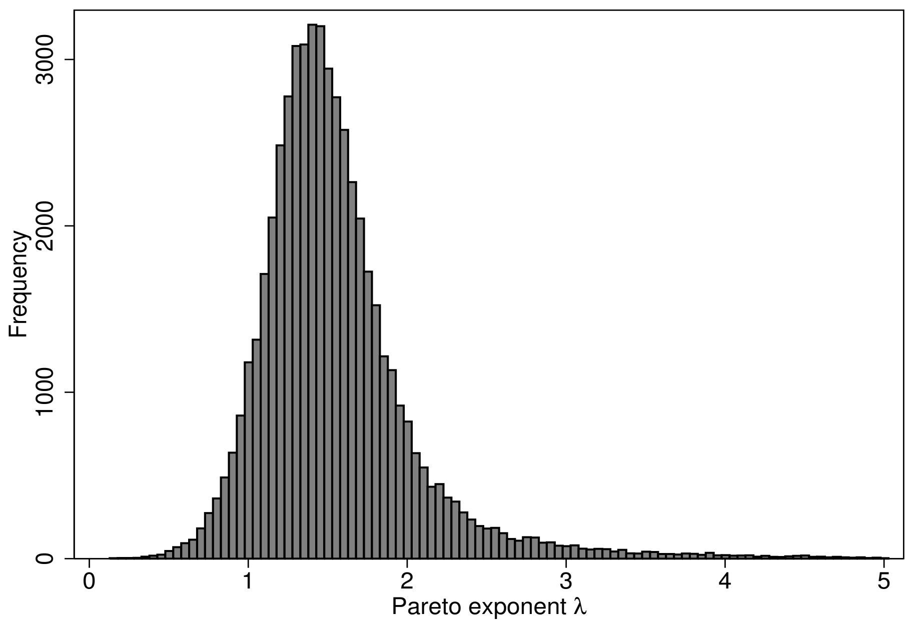
Notes: Calculations based on nationally representative household survey (*Susenas*) data collected in July 2005. The nonparametric regression curve and analytic confidence band is based on a local linear probability regression of an indicator for whether of a household member worked abroad from 2002-2005 on log landholdings under household control. The estimates employ a bandwidth of 0.4 and an Epanechnikov kernel. There are a total of 257,906 households in the data and 124,472 report controlling any landholdings at the time of enumeration. Both the mean estimate for migration probabilities in landless households and the nonparametric regression employ sampling weights. The top percentile of landholdings are trimmed from the figure for presentational purposes.

Figure 1.2: Migrants Drawn from the Middle of the Landholdings Distribution



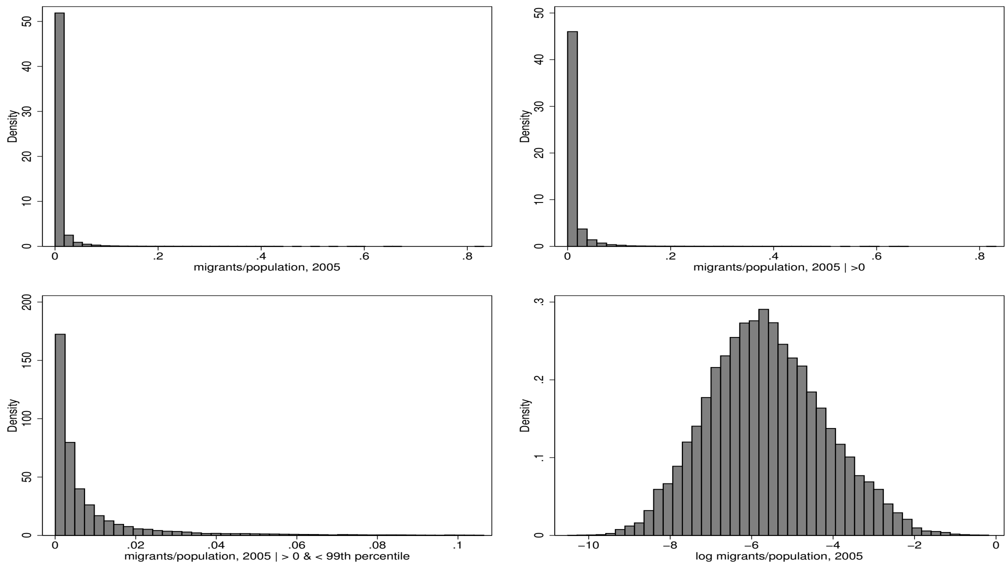
Notes: The figures report the log CCDF – log size observations for agricultural holdings for Indonesian households recorded in 16 districts chosen at random from the Agricultural Census of 2003. The graphs impose lower thresholds of $\underline{R} = 0.1$ in estimating the CCDF. The line constitutes the best linear fit from the log rank — log size regression. See Appendix 1.9.6 for further discussion.

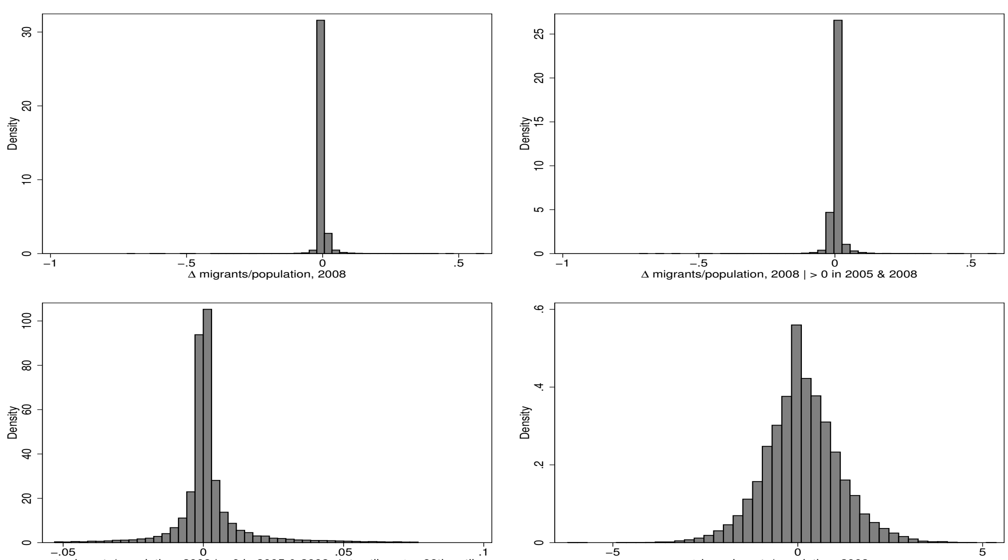
Figure 1.3: Pareto Linearity in Log–Log Plots



Notes: The Pareto distribution is given by $\lambda_v \underline{R}^{\lambda_v} R_{iv}^{-\lambda_v - 1}$. The figure shows the distribution of Gabaix and Ibragimov (2011) log rank minus 1/2 – log size OLS estimates of λ_v using the average log rank for a given log landholding size and imposing $\underline{R} = 0.1$ hectares. The estimates were calculated independently across 58,643 villages with at least 3 distinct total agricultural landholding sizes recorded in the Agricultural Census 2003. In the figure, the top 2 % of estimates are trimmed and bins are set to a width of 0.05.

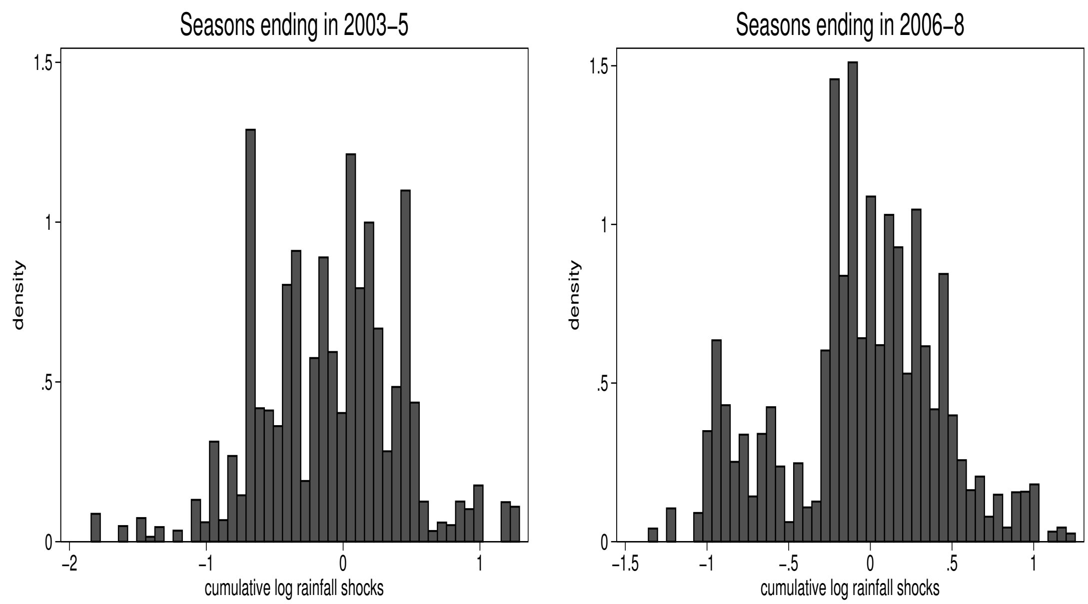
Figure 1.4: Distribution of Estimated Pareto Exponents λ_v





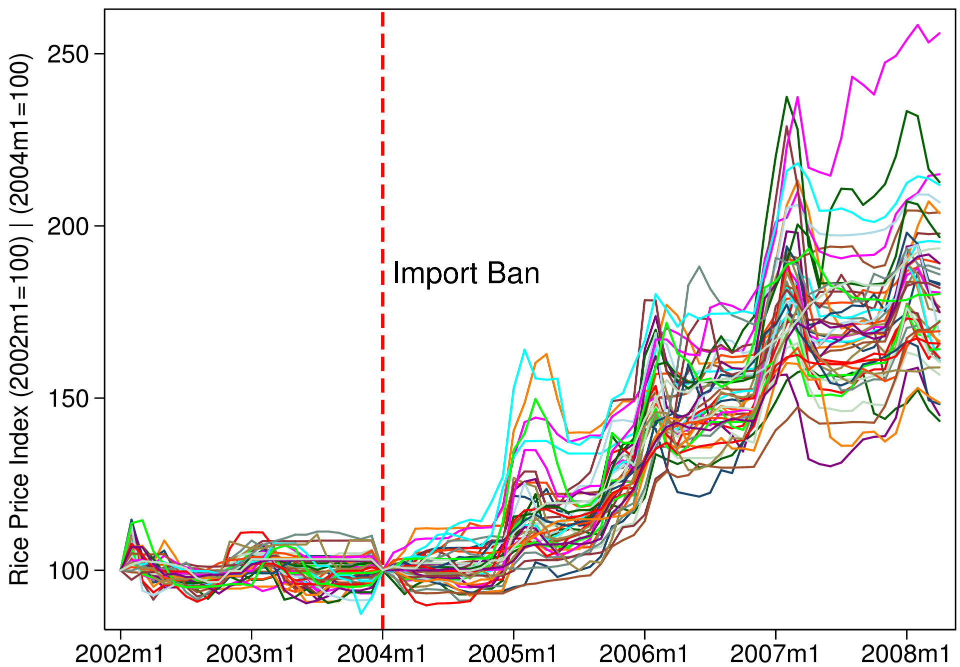
Notes: The top figure shows the distribution of migrants/population in 2005, and the bottom figure shows the distribution of the difference in migrants/population between 2008 and 2005.

Figure 1.5: Distribution of Migration Rates



Notes: The histograms show the spatial incidence of cumulative rainfall shocks over the growing seasons ending in 2003-5 and 2006-8. The shocks corresponding to each year are defined as the log difference between the given village's rainfall (measured at the district level) in the province-specific rice growing season and the long-run district-level mean rainfall excluding the given season from 1953-2008. Further details on the time series properties of rainfall can be found in Appendix 1.9.5.

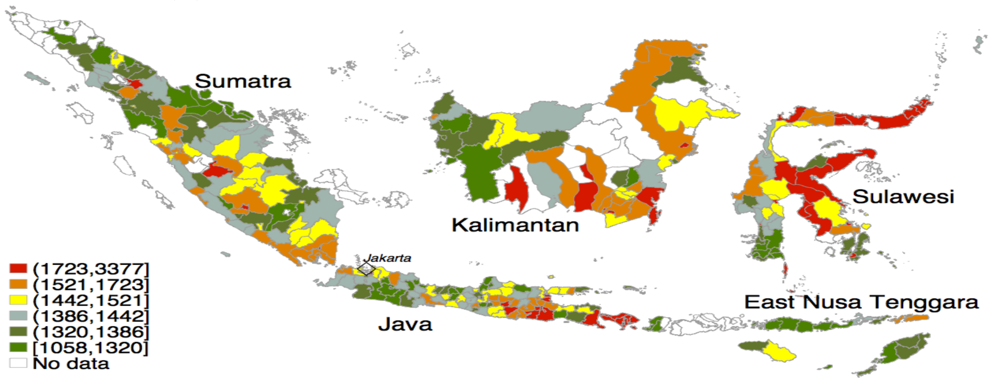
Figure 1.6: Rainfall Shocks Across the Indonesian Archipelago, 2002-2008



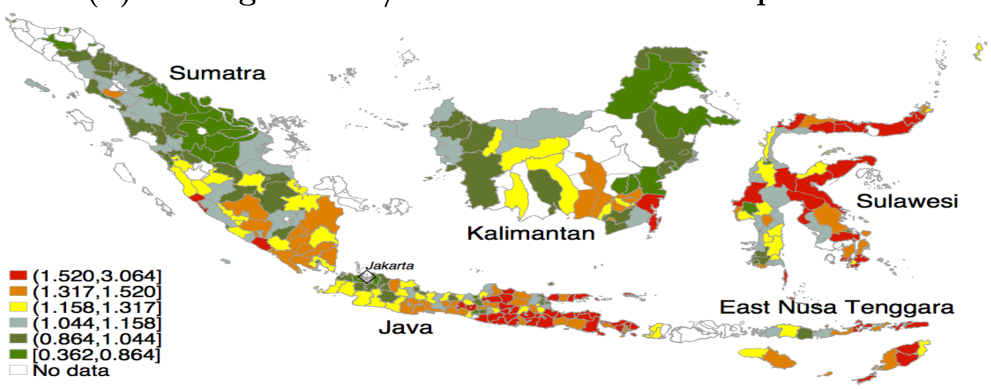
Notes: The index is initially normalized to equal 100 in January 2002. For the purposes of comparing before and after the ban, the bottom graph re-initializes and re-normalizes the index to equal 100 at the time of the import ban in January 2004. The rice price index is produced by the Central Bureau of Statistics for cities across the Indonesian archipelago based on prices collected in major markets within those cities. Though these estimates are based on consumer retail prices, I discuss evidence in Appendix 1.9.5 suggesting that retail, farmgate and wholesale prices move in lock-step from 2000-2008. The data were obtained from Wimanda (2009).

Figure 1.7: The Evolution of Rice Prices in Indonesian Cities, 2002-2008

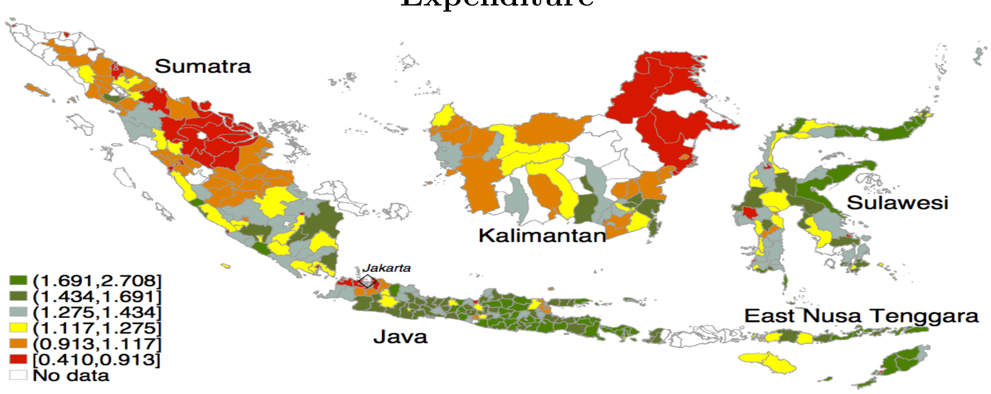
(a) Average Costs for Two-Year Contracts



(b) Average Costs / Annual Household Expenditure



(c) Average Two-Year Net Wages / Cumulative Household Expenditure



Notes: The colors correspond to sextiles. The village-specific costs (in 2006 USD) are recovered from the structural model for all villages in the baseline regression from column 14 of Table 1.6. To obtain the district averages, I weight each village's costs by its population in 2005. Estimates are missing for certain districts on account of villages in those districts being excluded from the two-step model as a result of missing data from one of the main datasets (see Appendix 1.9.8) or no households with wetland holdings in the district.

Figure 1.8: Estimated Village-Specific Migration Costs, by District

1.8 Tables

Table 1.1: International Labor Migration from Indonesian Villages

	Stocks, 2005			
	mean	*median*	*std. dev*	*max*
village population	3,216	2,095	4,123	78,986
number of emigrants	17	1	61	1,996
emigrants/population	0.006	0.0005	0.019	0.832
1(any emigrants abroad)	0.54	—	—	—
number of emigrants | emigrants > 0	31	8	80	1,996
emigrants/population | emigrants > 0	0.010	0.003	0.025	0.832
	Stocks, 2008			
	mean	*median*	*std. dev*	*max*
village population	3,377	2,187	4,330	82,215
number of emigrants	20	2	64	998
emigrants/population	0.007	0.0008	0.020	0.759
1(any emigrants abroad)	0.59	—	—	—
number of emigrants | emigrants > 0	35	9	81	998
emigrants/population | emigrants > 0	0.012	0.004	0.026	0.759
	Changes (Δ), 2005–2008			
	mean	*median*	*std. dev*	*max*
Δ number of emigrants	4	0	523	998
Δ emigrants/population	**0.110**	**0**	**1.918**	**59.2**
Δ number of emigrants | emigrants> 0	**6**	**1**	**73**	**995**
Δ emigrants/population | emigrants> 0	**0.143**	**0.016**	**2.551**	**59.2**
Δ log emigrants/population	**0.106**	**0.062**	**1.012**	**5.669**
	2005		**2008**	
share of Indonesian population from rural areas	0.59		0.59	
share of Indonesian migrants from rural areas	0.83		0.89	
total emigrants, all villages	1,113,244		1,349,540	

Notes: The statistics are calculated for all 65,966 villages matched in *Podes* 2005 and 2008. The qualitative patterns remain unchanged when restricting to villages with agricultural activities and assets recorded in the Agricultural Census of 2003. However, the mean and median stock migration rate figures are quantitatively larger when restricting to these non-urban villages. See Appendix 1.9.3 for a description of the determinants of rural area.

Table 1.2: Reduced Form Estimates

Estimator	OLS				Trimmed LAD or Semiparametric Tobit			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
rainfall shock	0.0011	0.0007	0.0011	0.0014	0.0083	0.0041	0.0083	0.0033
	(0.0002)***	(0.0006)	(0.0002)***	(0.0006)**	(0.0011)***	(0.0032)	(0.0011)***	(0.0047)
	[0.0011]	[0.0011]	[0.0011]	[0.0012]	[0.0047]*	[0.0050]	[0.0046]*	[0.0051]
rice price shock	0.0086	0.0086	0.0027	0.0023	0.0254	0.0253	-0.0179	-0.0185
	(0.0014)***	(0.0014)***	(0.0023)	(0.0023)	(0.0046)***	(0.0045)***	(0.0117)	(0.0135)
	[0.0058]	[0.0058]	[0.0062]	[0.0064]	[0.0172]	[0.0171]	[0.0172]	[0.0176]
rainfall shock × $\widehat{\lambda}_v$		0.0003		-0.0002		0.0027		0.0032
		(0.0004)		(0.0004)		(0.0021)		(0.0032)
		[0.0006]		[0.0006]		[0.0027]		[0.0033]
rice price shock × $\widehat{\lambda}_v$			0.0037	0.0040			0.0289	0.0293
			(0.0013)***	(0.0013)***			(0.0080)***	(0.0093)***
			[0.0023]*	[0.0026]			[0.0121]**	[0.0125]**
Observations	103,196	103,196	103,196	103,196	103,196	103,196	103,196	103,196
Years	2	2	2	2	2	2	2	2

Notes: This table reports estimates of $\boldsymbol{\theta}$ parameters in equation (1.12). The dependent variable in all specifications is the stock migration rate in village v in period t (migrants/population or M_{vt}/N_{vt}) the mean of which is 0.006 in the full sample of villages. *rainfall shock* is the cumulative log deviation from long-run mean rainfall in the growing seasons ending in 2006-2008 for t and 2002-2005 for $t-1$. *rice price shock* is the respective annualized log growth rate between 2005m4-2008m3 and 2002m1-2005m3. The estimated Pareto exponent $\widehat{\lambda}_v$ is obtained for total agricultural landholdings. The regressions also include an indicator for period 2 (not reported). In the two-period panel with fixed effects, the OLS specification is equivalent to first-differences. To interpret the coefficients in columns 5-8 as marginal effects, one simply multiplies the value for a given variable by the share of the sample with non-zero migration rates, 0.609. See Honoré (2008) for details. Standard errors in parentheses $(\cdot)$ are clustered at the village level and in brackets $[\cdot]$ are additionally robust to clustering within districts. A cluster/block bootstrap is used to obtain the latter for the semiparametric estimator.

Table 1.3: Validating the Theoretical Threshold Formulation for the Extensive Margin

Estimator:	SU-LPM		SNP-ML		Bivariate Probit	
	2008	2005	2008	2005	2008	2005
	(1)		(2)		(3)	
log max landholdings	0.023	0.031	0.056	0.102	0.088	0.109
	(0.005)***	(0.006)***	(0.011)***	(0.012)***	(0.020)***	(0.020)***
log min landholdings	-0.064	-0.057	-0.173	-0.179	-0.242	-0.222
	(0.012)***	(0.011)***	(0.017)***	(0.018)***	(0.043)***	(0.044)***
Number of villages	51,594	51,594	51,592	51,592	51,598	51,598

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The specifications are the same as in Table 1.4 below except that $\widehat{\lambda}_v$ and population size are not included as regressors since they theoretically determine the location of maximum and minimum landholdings. The minimum landholdings are calculated over all landholding sizes above $\underline{R} = 0.1$ Ha. The maximum landholdings include the extreme outliers, but the results are unchanged when considering the 2nd or 3rd largest landholding size instead.

Table 1.4: Extensive Margin First-Stage Estimates for Table 1.5

Estimator:	SU-LPM		SNP-ML		Bivariate Probit	
	2008	2005	2008	2005	2008	2005
	(1)		(2)		(3)	
Pareto exponent $\widehat{\lambda}_v$	-0.011	-0.016	-0.025	-0.051	-0.049	-0.069
	(0.005)**	(0.006)***	(0.013)*	(0.020)**	(0.021)**	(0.023)***
share households above 0.1 Ha	0.116	0.101	0.483	0.438	0.415	0.368
	(0.026)***	(0.027)***	(0.048)***	(0.048)***	(0.094)***	(0.095)***
log village population, s	0.081	0.074	0.309	0.292	0.304	0.277
	(0.006)***	(0.006)***	(0.022)***	(0.026)***	(0.020)***	(0.021)***
log district population less v, s	0.095	0.091	0.353	0.295	0.316	0.303
	(0.036)***	(0.033)***	(0.021)***	(0.024)***	(0.109)***	(0.101)***
log district area less v	-0.047	-0.053	-0.245	-0.241	-0.156	-0.178
	(0.018)***	(0.017)***	(0.022)***	(0.024)***	(0.056)***	(0.051)***
log number of villages in district	0.002	0.019	0.070	0.114	0.021	0.073
	(0.047)	(0.041)	(0.027)***	(0.030)***	(0.135)	(0.116)
rice price shock, $s-1$	0.041	0.139	1.008	4.468	0.193	0.499
	(0.416)	(0.408)	(0.415)**	(0.618)***	(1.364)	(1.217)
rainfall shock, $s-1$	0.027	0.034	0.258	0.301	0.078	0.108
	(0.026)	(0.025)	(0.033)***	(0.036)***	(0.095)	(0.084)
log distance to subdistrict capital	-0.022	-0.023	-0.048	-0.058	-0.087	-0.087
	(0.005)***	(0.005)***	(0.011)***	(0.011)***	(0.016)***	(0.015)***
log distance to district capital	-0.015	-0.015	0.006	0.004	-0.060	-0.062
	(0.009)	(0.009)	(0.013)	(0.013)	(0.032)*	(0.029)**
log distance to nearest emigration center	-0.026	-0.021	-0.410	-0.315	-0.021	-0.016
	(0.029)	(0.029)	(0.064)***	(0.055)***	(0.087)	(0.080)
Number of villages	51,592	51,592	51593	51593	51,592	51,592

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The table reports estimates of the extensive margin in equations (1.10). The dependent variable equals one if the village reports any residents working abroad at time s. The estimated Pareto exponent $\widehat{\lambda}_v$ is for total agricultural landholdings; results are similar for other land-holdings measures. Additional covariates not reported here include: log distance to the subdistrict capital and nearest emigration center, the share of households with landholdings above 0.1 hectares, the share of wetland in total agricultural land in the village, an indicator for whether the government classifies the village as urban, Muslim population share, ethnic Chinese population share, ethnic Arab population share, and an indicator for whether the village is accessible by motorized land transport. SU-LPM connotes seemingly unrelated linear probability models and estimates the LPM models for 2005 and 2008 jointly through a feasible generalized least squares procedure. SNP-ML connotes a semi-nonparametric maximum likelihood procedure due to Gallant and Nychka (1987). Standard errors are clustered at the district level (using a block bootstrap procedure for the SU-LPM estimates) for all specifications except the SNP-ML which are merely heteroskedasticity-robust and hence should be heavily discounted.

Table 1.5: Two-Step Estimates of Flow Migration Rates

Correction Procedure	None	Semiparametric		Parametric
1st Stage Estimator	None	SU-LPM	SNP-ML	BiProbit
Landholdings measure	**Total Agricultural Landholdings**			
	(1)	(2)	(3)	(4)
Pareto exponent $\widehat{\lambda}_v$	-0.002	0.039	0.039	0.038
	(0.018)	(0.017)*	(0.018)**	(0.018)*
Δ rice price shock	**-0.092**	**0.409**	**3.504**	**0.283**
	(0.438)	**(0.448)**	**(0.850)*** **	**(0.426)**
Δ rainfall shock	**0.098**	**0.415**	**0.572**	**0.296**
	(0.127)	**(0.133)*** **	**(0.156)*** **	**(0.128)** **
Number of Villages	26,529	26,527	26,527	26,527
R^2	0.021	0.036	0.032	0.032
Landholdings measure	**Total Wetland Holdings**			
	(5)	(6)	(7)	(8)
Pareto exponent $\widehat{\lambda}_v$	-0.018	0.072	0.002	0.050
	(0.014)	(0.018)***	(0.015)	(0.017)***
Δ rice price shock	**-0.111**	**1.287**	**3.631**	**0.625**
	(0.443)	**(0.502)***	**(1.030)*** **	**(0.451)**
Δ rainfall shock	**0.048**	**0.305**	**0.288**	**0.212**
	(0.132)	**(0.134)** **	**(0.145)** **	**(0.130)***
Number of Villages	24,539	24,537	24,537	24,537
R^2	0.019	0.025	0.025	0.030
Landholdings measure	**Total Paddy Area Planted, 2002-3**			
	(9)	(10)	(11)	(12)
Pareto exponent $\widehat{\lambda}_v$	0.011	0.043	0.034	0.048
	(0.016)	(0.016)**	(0.014)**	(0.017)**
Δ rice price shock	**-0.134**	**0.919**	**2.034**	**0.457**
	(0.433)	**(0.487)**	**(0.307)** **	**(0.447)**
Δ rainfall shock	**0.086**	**0.390**	**0.366**	**0.253**
	(0.133)	**(0.139)** **	**(0.067)*** **	**(0.136)** **
Number of villages	24,855	24,855	24,855	24,855
R^2	0.021	0.036	0.032	0.032

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The table reports estimates of equation (1.13) in the text. The dependent variable in all specifications is $\Delta \ln(M_{v,t+1}/N_{v,t+1})$ and has mean 0.11. Δ rainfall shock is the difference in cumulative log deviations from long-run mean rainfall between the growing seasons ending in 2006-2008 and 2002-2005. Δ rice price shock is the difference in annualized log growth rates between 2005m4-2008m3 and 2002m1-2005m3. Each of the three panels uses the given landholdings measure to estimate λ_v. Additional covariates in all specifications but not reported here include those all those in Table 1.4. Column 4/8/12 additionally includes the bivariate Mills ratios given in Appendix 1.9.2. Columns 2-3/6-7/10-11 include a 3rd degree polynomial in the propensity scores for the extensive margin in 2005 and 2008 based on the given estimator listed at the top of each column. These correction terms are jointly statistically significant in all specifications. The excluded instruments in this table and all subsequent tables are as reported in Table 1.4. Standard errors are clustered at the district level, and the significance levels are based on a block bootstrap-t procedure.

Table 1.6: Landholdings Heterogeneity, Agricultural Income Shocks, and Flow Migration Rates

Correction Procedure	None	Semipar.	Param.	None	Semipar.	Param.	None	Semipar.	Param.
1st Stage Estimator	None	SU-LPM	BiProbit	None	SU-LPM	BiProbit	None	SU-LPM	BiProbit
Landholdings measure	**Total Agri. Landholdings**			**Wetland Holdings**			**Paddy Planted, 2002-3**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Δ rainfall shock	0.083	0.184	0.132	0.080	0.162	0.069	-0.072	0.113	0.045
	(0.164)	(0.167)	(0.158)	(0.163)	(0.167)	(0.166)	(0.166)	(0.171)	(0.170)
Pareto exponent $\widehat{\lambda}_v$	-0.001	0.017	0.036	-0.020	0.073	0.052	0.011	0.027	0.044
	(0.018)	(0.018)	(0.018)*	(0.015)	(0.018)***	(0.017)***	(0.016)	(0.017)	(0.017)**
$\widehat{\boldsymbol{\lambda}}_v \times \boldsymbol{\Delta}$ **rainfall shock**	**0.011**	**0.147**	**0.108**	**-0.019**	**0.082**	**0.085**	**0.098**	**0.165**	**0.126**
	(0.070)	**(0.072)****	**(0.067)***	**(0.049)**	**(0.049)**	**(0.058)**	**(0.054)***	**(0.059)*****	**(0.057)***
Landholdings measure	**Total Agri. Landholdings**			**Wetland Holdings**			**Paddy Planted, 2002-3**		
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Pareto exponent $\widehat{\lambda}_v$	0.038	-0.007	-0.019	0.008	-0.113	-0.070	0.028	-0.083	-0.084
	(0.038)	(0.035)	(0.034)	(0.022)	(0.030)***	(0.030)***	(0.031)	(0.046)*	(0.042)***
Δ rainfall shock	0.115	0.225	0.188	0.071	0.262	0.135	-0.064	0.167	0.110
	(0.164)	(0.169)	(0.162)	(0.162)	(0.168)	(0.161)	(0.168)	(0.176)	(0.174)
$\widehat{\boldsymbol{\lambda}}_v \times \boldsymbol{\Delta}$ **rainfall shock**	**-0.013**	**0.119**	**0.073**	**-0.018**	**0.028**	**0.048**	**0.091**	**0.140**	**0.087**
	(0.069)	**(0.074)****	**(0.070)**	**(0.047)**	**(0.052)**	**(0.051)**	**(0.055)***	**(0.065)****	**(0.061)**
Δ price shock	0.551	-0.016	-0.616	0.453	-2.234	-1.503	0.186	-1.031	-1.750
	(0.732)	(0.688)	(0.686)	(0.603)	(0.709)*	(0.688)*	(0.686)	(0.822)	(0.776)**
$\widehat{\boldsymbol{\lambda}}_v \times \boldsymbol{\Delta}$ **price shock**	**-0.416**	**0.267**	**0.586**	**-0.310**	**1.913**	**1.155**	**-0.192**	**1.116**	**1.314**
	(0.352)	**(0.329)**	**(0.335)***	**(0.232)**	**(0.335)*****	**(0.327)*****	**(0.291)**	**(0.423)*****	**(0.400)*****
Number of villages	26,529	26,527	26,527	24,540	24,537	24,537	24,856	24,855	24,855

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The table reports estimates of equation (1.14) and (1.15) in the text. The dependent variable in all specifications is $\Delta \ln(M_{v,t+1}/N_{v,t+1})$ and has mean 0.11. Sample sizes are identical across sub-columns within the super-column, as reported at the bottom of the table. Additional covariates in all specifications but not reported here include all those in (the notes to) Table 1.5. The semiparametric (Semipar.) estimator includes a 3rd degree polynomial in the propensity scores for the extensive margin in 2005 and 2008 based on a SU-LPM estimator for the extensive margin, and the parametric (Param.) correction procedure is based on the Poirier procedure. The correction terms are jointly statistically significant in all specifications. Standard errors are clustered at the district level, and the significance levels are based on a block bootstrap-t procedure.

Table 1.7: Landholdings Heterogeneity, Income Shocks, and Migration Flows

Correction Procedure	None	Semipar.	Param.	None	Semipar.	Param.	None	Semipar.	Param.
1st Stage Estimator	None	SU-LPM	BiProbit	None	SU-LPM	BiProbit	None	SU-LPM	BiProbit
Landholdings measure	**Total Agri. Landholdings**			**Wetland Holdings**			**Paddy Planted, 2002-3**		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Δ rainfall shock	0.085	0.740	0.668	0.141	0.561	0.450	-0.002	0.518	0.445
	(0.188)	(0.179)***	(0.164)***	(0.167)	(0.217)***	(0.200)***	(0.174)	(0.210)***	(0.205)***
Pareto exponent $\widehat{\lambda}_v$	-0.003	0.015	0.034	-0.021	0.046	0.037	0.009	0.021	0.037
	(0.018)	(0.018)	(0.018)*	(0.015)	(0.018)**	(0.017)***	(0.016)	(0.017)	(0.017)**
$\times$ Δ rainfall shock	**-0.002**	**0.062**	**0.030**	**-0.018**	**0.042**	**0.031**	**0.103**	**0.110**	**0.071**
	(0.067)	**(0.068)**	**(0.064)**	**(0.049)**	**(0.055)**	**(0.057)**	**(0.051)****	**(0.058)***	**(0.060)**
share households < 0.1 Ha	-0.026	-0.026	-0.019	-0.019	0.039	0.022	0.003	-0.048	-0.029
	(0.048)	(0.047)	(0.047)	(0.047)	(0.047)	(0.047)	(0.044)	(0.045)	(0.045)
$\times$ Δ rainfall shock	**-0.097**	**-0.789**	**-0.682**	**0.067**	**-0.486**	**-0.330**	**0.101**	**-0.487**	**-0.366**
	(0.144)	**(0.133)*****	**(0.138)*****	**(0.135)**	**(0.151)*****	**(0.141)*****	**(0.131)**	**(0.148)*****	**(0.142)*****
share pop. paid agricultural labor	-0.129	-0.451	-0.316	-0.134	-0.337	-0.250	-0.149	-0.389	-0.291
	(0.068)*	(0.087)***	(0.075)***	(0.072)*	(0.090)***	(0.081)***	(0.071)**	(0.090)***	(0.081)***
$\times$ Δ rainfall shock	**-0.363**	**-1.124**	**-0.922**	**-0.383**	**-0.792**	**-0.709**	**-0.379**	**-0.997**	**-0.849**
	(0.291)	**(0.334)*****	**(0.306)*****	**(0.299)**	**(0.347)****	**(0.316)*****	**(0.300)**	**(0.342)*****	**(0.320)*****
Landholdings measure	**Total Agri. Landholdings**			**Wetland Holdings**			**Paddy Planted, 2002-3**		
	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Δ price shock	0.105	0.982	0.510	-0.067	0.260	0.159	-0.473	0.944	-0.122
	(0.858)	(0.819)	(0.814)	(0.802)	(0.947)	(0.868)	(0.841)	(1.042)	(0.925)
Δ rainfall shock	0.196	0.790	0.713	0.033	0.668	0.491	-0.126	0.527	0.435
	(0.159)	(0.180)***	(0.166)***	(0.178)	(0.208)***	(0.195)**	(0.168)	(0.206)**	(0.201)**
Pareto exponent $\widehat{\lambda}_v$	0.033	0.016	0.011	0.006	-0.073	-0.040	0.024	-0.050	-0.038
	(0.039)	(0.035)	(0.035)	(0.023)	(0.029)**	(0.028)	(0.031)	(0.041)	(0.038)
$\times$ Δ price shock	**-0.381**	**-0.026**	**0.236**	**-0.303**	**1.162**	**0.675**	**-0.170**	**0.656**	**0.757**
	(0.352)	**(0.322)**	**(0.336)**	**(0.232)**	**(0.311)*****	**(0.291)****	**(0.289)**	**(0.374)***	**(0.360)****
$\times$ Δ rainfall shock	**-0.024**	**0.025**	**-0.009**	**-0.016**	**-0.036**	**-0.004**	**0.099**	**0.091**	**0.045**
	(0.065)	**(0.068)**	**(0.067)**	**(0.047)**	**(0.051)**	**(0.051)**	**(0.052)***	**(0.058)**	**(0.062)**
share households < 0.1 Ha	-0.193	-0.831	-0.527	-0.074	0.281	0.181	-0.077	0.233	0.137
	(0.122)	(0.167)***	(0.141)***	(0.081)	(0.091)***	(0.078)**	(0.076)	(0.086)***	(0.076)*
$\times$ Δ price shock	**0.704**	**-2.251**	**-1.912**	**0.574**	**-2.820**	**-1.842**	**0.831**	**-3.076**	**-1.756**
	(0.799)	**(0.795)****	**(0.774)*****	**(0.739)**	**(0.870)*****	**(0.760)****	**(0.711)**	**(0.892)*****	**(0.777)****
$\times$ Δ rainfall shock	**-0.359**	**-1.151**	**-0.853**	**0.108**	**-0.458**	**-0.327**	**0.140**	**-0.456**	**-0.319**
	(0.297)	**(0.350)*****	**(0.329)*****	**(0.141)**	**(0.152)*****	**(0.143)****	**(0.133)**	**(0.150)*****	**(0.141)****
share pop. paid agricultural labor	-0.094	0.175	0.156	-0.227	-0.725	-0.549	-0.243	-0.760	-0.499
	(0.081)	(0.079)**	(0.076)**	(0.121)*	(0.196)***	(0.155)***	(0.120)**	(0.185)***	(0.143)***
$\times$ Δ price shock	**0.594**	**3.767**	**2.167**	**0.881**	**3.903**	**2.971**	**0.896**	**3.568**	**2.235**
	(1.104)	**(1.485)***	**(1.275)**	**(1.107)**	**(1.718)****	**(1.382)****	**(1.091)**	**(1.603)****	**(1.295)***
$\times$ Δ rainfall shock	**-0.071**	**-0.801**	**-0.693**	**-0.363**	**-0.865**	**-0.693**	**-0.360**	**-1.051**	**-0.765**
	(0.151)	**(0.138)*****	**(0.146)*****	**(0.304)**	**(0.356)****	**(0.342)****	**(0.305)**	**(0.356)*****	**(0.345)****
Number of villages	26,529	26,527	26,527	24,540	24,537	24,537	24,856	24,855	24,855

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The table augments the specifications in Table 1.6 with interactions of the income shock terms and other measures of local agricultural activity. The dependent variable in all specifications is $\Delta \ln(M_{v,t+1}/N_{v,t+1})$ and has mean 0.11. Sample sizes are identical across sub-columns within the super-column, as reported at the bottom of the table. Additional covariates in all specifications but not reported here include all those in (the notes to) Table 1.5. The semiparametric (Semipar.) estimator includes a 3rd degree polynomial in the propensity scores for the extensive margin in 2005 and 2008 based on a SU-LPM estimator for the extensive margin, and the parametric (Param.) correction procedure is based on the Poirier procedure. The correction terms are jointly statistically significant in all specifications. Standard errors are clustered at the district level, and the significance levels are based on a block bootstrap-t procedure.

Table 1.8: Further Evidence of Liquidity Constraints

Correction Procedure	None	Semipar.	Param.	None	Semipar.	Param.
1st Stage Estimator	None	SU-LPM	BiProbit	None	SU-LPM	BiProbit
	z := bank presence in village or subdistrict, 2005			**z := log mean household expenditures/capita, 2000**		
	(1)	(2)	(3)	(4)	(5)	(6)
$z \times \Delta$ rainfall shock	**-0.031**	**-0.226**	**-0.162**	**-0.145**	**-0.909**	**-0.841**
	(0.092)	**(0.087)****	**(0.085)****	**(0.160)**	**(0.163)*****	**(0.150)*****
Δ rainfall shock	0.122	0.572	0.419	1.788	10.897	10.037
	(0.145)	(0.155)***	(0.147)***	(1.844)	(1.912)***	(1.742)***
z	-0.012	-0.114	-0.082	-0.043	-0.014	-0.062
	(0.025)	(0.026)***	(0.025)***	(0.056)	(0.059)	(0.057)
Number of Villages	26,529	26,527	26,527	26,129	26,127	26,127
	z := technical irrigation in village, 2005			**z := log total rice output (tons) per Ha in village, 2001**		
	(7)	(8)	(9)	(10)	(11)	(12)
$z \times \Delta$ rainfall shock	**-0.078**	**-0.256**	**-0.187**	**-0.020**	**-0.126**	**-0.097**
	(0.066)	**(0.068)*****	**(0.068)*****	**(0.031)**	**(0.034)*****	**(0.031)*****
Δ rainfall shock	0.149	0.527	0.396	0.132	0.471	0.375
	(0.130)	(0.137)***	(0.132)***	(0.134)	(0.149)***	(0.140)***
z	0.011	-0.014	0.007	-0.013	-0.014	-0.015
	(0.019)	(0.024)	(0.022)	(0.009)	(0.010)	(0.009)
Number of villages	26,129	26,127	26,127	26,529	26,527	26,527

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The table reports estimates of equation (1.16) in the text. The dependent variable in all specifications is $\Delta \ln(M_{v,t+1}/N_{v,t+1})$ and has mean 0.11. Additional covariates in all specifications but not reported here include all those in (the notes to) Table 1.5. Bank presence equals one if any banks were located in village or subdistrict in 2005 and zero otherwise; log mean household expenditures/capita in 2000 obtained from Poverty Map estimates by Suryahadi et al. (2005); technical irrigation equals one if in 2005 the village has any land irrigated by a technical system not reliant on rainfall; total rice output per Ha is based on village-level records from the 2001 growing seasons as reported in *Podes* 2002. See Appendix 1.9.2 for details on the semiparametric (Semipar.) and parametric (Param.) correction procedures. Standard errors are clustered at the district level, and the significance levels are based on a block bootstrap-t procedure.

Table 1.9: Relaxing Sets of Exclusion Restrictions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pareto exponent $\hat{\lambda}_v$	0.072	0.073	-0.113	0.080	0.082	-0.111	0.090	0.091	-0.109
	(0.018)***	(0.018)***	(0.030)***	(0.018)***	(0.018)***	(0.030)**	(0.019)***	(0.019)***	(0.031)**
Δ rainfall shock	**1.287**	**1.304**	**-2.234**	**1.299**	**1.317**	**-2.343**	**1.376**	**1.397**	**-2.381**
	(0.502)*	**(0.503)***	**(0.709)***	**(0.513)**	**(0.515)**	**(0.724)**	**(0.507)**	**(0.509)**	**(0.734)**
Δ price shock	**0.305**	**0.162**	**0.262**	**0.328**	**0.150**	**0.260**	**0.346**	**0.165**	**0.277**
	(0.134)**	**(0.167)**	**(0.168)**	**(0.129)****	**(0.165)**	**(0.166)**	**(0.128)****	**(0.166)**	**(0.167)**
$\hat{\lambda}_v \times \Delta$ rainfall shock		**0.082**	**0.028**		**0.103**	**0.040**		**0.105**	**0.040**
		(0.049)	**(0.052)**		**(0.048)***	**(0.052)**		**(0.048)***	**(0.053)**
$\hat{\lambda}_v \times \Delta$ price shock			**1.913**			**1.973**			**2.031**
			(0.335)***			**(0.334)*****			**(0.344)*****
log # villages in district				-0.073	-0.076	-0.065	-0.127	-0.131	-0.114
				(0.025)*	(0.025)**	(0.025)*	(0.034)	(0.034)*	(0.033)
log district area less v							0.057	0.058	0.052
							(0.025)	(0.025)	(0.024)

	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)
Pareto exponent $\hat{\lambda}_v$	0.064	0.065	-0.108	0.064	0.065	-0.108	0.067	0.069	-0.108
	(0.018)**	(0.019)**	(0.030)***	(0.018)**	(0.019)**	(0.030)**	(0.019)**	(0.019)**	(0.030)**
Δ rainfall shock	**1.163**	**1.183**	**-2.131**	**1.155**	**1.173**	**-2.134**	**1.160**	**1.178**	**-2.168**
	(0.518)	**(0.519)**	**(0.726)**	**(0.521)**	**(0.522)**	**(0.726)**	**(0.521)**	**(0.523)**	**(0.724)***
Δ price shock	**0.271**	**0.123**	**0.217**	**0.271**	**0.120**	**0.215**	**0.276**	**0.116**	**0.216**
	(0.132)*	**(0.168)**	**(0.169)**	**(0.132)**	**(0.168)**	**(0.169)**	**(0.131)**	**(0.168)**	**(0.169)**
$\hat{\lambda}_v \times \Delta$ rainfall shock		**0.086**	**0.034**		**0.087**	**0.035**		**0.093**	**0.036**
		(0.050)*	**(0.052)**		**(0.049)**	**(0.052)**		**(0.049)**	**(0.052)**
$\hat{\lambda}_v \times \Delta$ price shock			**1.787**			**1.785**			**1.805**
			(0.337)***			**(0.338)*****			**(0.337)*****
log village pop., t	0.089	0.089	0.086	0.090	0.090	0.087	0.076	0.073	0.078
	(0.028)*	(0.028)*	(0.027)	(0.027)	(0.027)	(0.027)	(0.035)	(0.036)	(0.034)
log district pop. less v, t							-0.025	-0.030	-0.016
							(0.038)	(0.038)	(0.038)
log district area less v				-0.006	-0.006	-0.005			
				(0.019)	(0.019)	(0.018)			
Number of villages	24,537	24,537	24,537	24,537	24,537	24,537	24,537	24,537	24,537

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The table reports estimates of the key elasticity parameters sequentially relaxing one or two of the four baseline exclusion restrictions. The estimates in columns 1-3 correspond to baseline results for $\widehat{\lambda}_v$ obtained for wetland holdings. All estimates are based on the Das et al. (2003) semiparametric correction procedure and the measure of $\widehat{\lambda}_v$ for wetland holdings. The results are similar for parametric Poirier (1980) correction procedure and others types of landholdings. The dependent variable in all specifications is $\Delta \ln(M_{v,t+1}/N_{v,t+1})$ and has mean 0.11. Additional covariates in all specifications but not reported here include all those in Table 1.5. The correction terms are jointly statistically significant in all specifications. Standard errors are clustered at the district level, and significance levels are based on a block bootstrap-t procedure. Sample sizes are identical across sub-columns within the super-column, as reported at the bottom of the table.

Table 1.10: Agricultural Income Shocks and Migration *Choice*—Micro Data

	(1)	(2)	(3)	(4)	(5)
rainfall shock, t	0.200	0.212	1.109	1.712	1.510
	(0.115)*	(0.112)*	(0.578)*	(0.605)***	(0.696)**
rainfall shock, t × landholdings (Ha)			-0.267	-3.844	
			(0.580)	(1.416)***	
rainfall shock, t × landholdings (Ha) squared				1.583	
				(0.647)**	
price shock, t		0.762	3.752	3.892	
		(0.340)**	(1.687)**	(1.729)**	
price shock, t × landholdings (Ha)			-0.581	-3.128	
			(1.176)	(1.875)*	
price shock, t × landholdings (Ha) squared				0.802	
				(0.513)	
rainfall shock, t × bank in village					-3.796
					(1.381)***
Average Marginal Effect (AME)	Yes	Yes	No	No	No
Point Estimate	No	No	Yes	Yes	Yes
Observations	1,902	1,380	1,380	1,380	1,902
Years	6	5	5	5	6

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The table reports conditional fixed effects logit of equation (1.17) for whether a household has any migrants departing in year $t+1$. The sample comprises a balanced panel of households with any migrants over the period 2000-2006 as recorded in the nationally representative *Susenas* household survey conducted in mid-2006. The rainfall shock in period t is defined as the log deviation of the current season's rainfall from the long-run local mean. The price shock in period t is defined as the log difference in the local rice price at the end of period t and $t-1$. All columns include year fixed effects. Standard errors are clustered at the district level as in the village-level regressions.

Table 1.11: Comparing Elasticities Based on Micro and Macro Data

Elasticity Summary Statistic⟶	mean	std. dev.	25th pctile	median	75th pctile	correlation w/ village-level reg.
Δprice shock						
Aggregating Micro Data AMEs (linear R_{iv})	0.771	0.018	0.766	0.775	0.782	0.898
Aggregating Micro Data AMEs (quadratic R_{iv})	0.698	0.030	0.683	0.702	0.718	0.757
Village-Level Regression	0.747	0.473	0.440	0.680	0.974	—
Δrainfall shock						
Aggregating Micro Data AMEs (linear R_{iv})	0.223	0.006	0.221	0.225	0.227	0.794
Aggregating Micro Data AMEs (quadratic R_{iv})	0.226	0.018	0.214	0.226	0.238	0.977
Village-Level Regression	0.389	0.059	0.351	0.381	0.417	—

Notes: This table reports elasticities of flow migration rates with respect to rainfall and rice price shocks based on two approaches. *Aggregating Micro Data AMEs* elasticities are computed based on columns 3-4 in Table 1.10 as the sum of nationally representative average marginal effects (AMEs) of shocks at landholding sizes $\in \{0.1, 0.2, \ldots, 2.5\}$ weighted by the share of households in the village falling within each of the given size ranges as implied by the estimated Pareto dispersion parameter $\widehat{\lambda}_v$. *Village-Level Regression* elasticities are based on the estimates of θ parameters using village-level data in column 5 of Table 1.7. All elasticities are based on λ_v estimated for total wetland holdings and are restricted to villages in the second-step sample (i.e., those with any migrants in 2005 and 2008).

Table 1.12: Summary Statistics on Estimated Village-Specific Migration Costs (in USD)

	mean	std. dev.	min	25th pctile	median	75th pctile	max	# of villages
all villages	1,485	430	108	1,295	1,373	1,469	5,011	42,063
villages on … island(s)								
Sumatra	1,432	252	469	1,343	1,390	1,457	4,385	9,711
Java/Bali	1,494	499	865	1,279	1,339	1,453	5,011	22,889
East/West Nusa Tenggara	1,288	178	108	1,159	1,271	1,344	3,815	1,538
Kalimantan	1,525	310	1,140	1,347	1,476	1,535	3,810	3,726
Sulawesi/Maluku	1,591	462	1,094	1,326	1,420	1,685	4,191	4,199

Notes: The village-specific costs (in, roughly, 2006 USD) are recovered from the structural equation (1.4) and are reported for all villages in the baseline regression from column 5 of Table 1.7. Details on the estimation procedure and auxiliary data inputs can be found in the text and in Appendix 1.9.4.

1.9 Appendix

1.9.1 Theoretical Results

In this section, I provide proofs of several results in Section 1.3. Equations (1.3)-(1.6) in Section 1.3.2 can be obtained by integrating over R_{iv} in equation (1.2) with (i) $\tau_{vj} > 0$ and $R_L \geq \underline{R}$ or (ii) $R_L < \underline{R}$ ($\Longleftarrow \tau_{vj} = 0$). First, consider the following expressions for the thresholds within which migration is both feasible and profitable in period t (from the perspective of $t-1$ decision-makers required to pay fixed upfront costs in that period)

$$R_{L,t-1} = \left(\frac{\tau_{vj} C_{vjt}}{p_{v,t-1}(\overline{\sigma}_v + a_{v,t-1}) K_v^\theta S_v^\phi} \right)^{1/\beta} \quad ; \quad R_{U,t-1} = \left(\frac{\omega_{vjt} - C_{vjt}}{\alpha_v p_{v,t-1} \overline{\sigma}_v K_v^\theta S_v^\phi} \right)^{1/\beta}, \tag{1.18}$$

where I have assumed that $\mathbb{E}_{t-1}[p_{vt}\sigma_{vt}] = \alpha_v p_{v,t-1}\overline{\sigma}_v$—a simplification hinges on the fact that $\text{cov}_{t-1}(p_{vt}\sigma_{vt}) = 0$, i.e. households cannot forecast the relationship between rainfall and prices next period. This does not imply that past rainfall has no effect on contemporaneous prices. Rather, $a_{v,t-k}$ for $k > 0$ are elements of the error term $\sum_{q=0}^{Q} \theta_q e_{v,t-q}$ in the $ARMA(1,Q)$ expression for rice prices. Thus, past output has a direct effect on current prices.

The expression for the migration rate in equation (1.3) is derived by integrating over all landholdings
$R_{iv} \in [R_{L,t-1}, R_{U,t-1}]$ in village v (maintaining the innocuous normalization $\underline{R} = 1$ Ha)

$$\mathbb{P}(R_{L,t-1} \leq R_{iv} \leq R_{U,t-1}) = \frac{M_{vt}}{N_{vt}} = \int_{R_{L,t-1}}^{R_{U,t-1}} \lambda_v R_{iv}^{-\lambda_v - 1} \mathrm{d}R_{iv} = R_{L,t-1}^{-\lambda_v} - R_{U,t-1}^{-\lambda_v}. \tag{1.19}$$

Figure 1.9 shows the stock migration rate M_{vt}/N_{vt}—the area below the solid density between R_L and R_U—for a given triplet (R_L, R_U, λ_v). Replacing the expressions for R_L and R_U with those in equation (1.18) and taking the difference between $t+1$ and t, we obtain equation (1.3). Taking logarithms and differences delivers equation (1.4).

On the other hand, if liquidity constraints are not binding and $R_L < \underline{R}$ ($\Longleftarrow \tau_{vj} = 0$), then

$$\mathbb{P}(1 \leq R_{iv} \leq R_{U,t-1}) = \frac{M_{vt}}{N_{vt}} = \int_1^{R_{U,t-1}} \lambda_v R_{iv}^{-\lambda_v - 1} \mathrm{d}R_{iv} = 1 - R_{U,t-1}^{-\lambda_v}. \tag{1.20}$$

Similarly substituting for $R_{U,t-1}$ and taking (log) differences delivers equations (1.5) and (1.6). Recall that, by definition, the expressions for the intensive margin in (1.19) and (1.20) must be greater than zero.

Proposition 1

The proofs in the presence of liquidity constraints follow immediately from differentiation of equation (1.4). Letting $\Delta \ln(M_{v,t+1}/N_{v,t+1}) \equiv \Delta \widehat{M}_{v,t+1}$,

$$\frac{\partial \Delta \widehat{M}_{v,t+1}}{\partial \Delta \ln p_{vt}} = \frac{\lambda_v}{\beta} > 0; \quad \frac{\partial \Delta \widehat{M}_{v,t+1}}{\partial a_{vt}} = \frac{\frac{\lambda_v}{\beta}(\overline{\sigma}_v + a_{vt})^{\lambda_v/\beta - 1}(\tau_{vj} C_{vj,t+1})^{-\lambda_v/\beta}}{\left(\frac{\overline{\sigma}_v + a_{vt}}{\tau_{vj} C_{vj,t+1}}\right)^{\lambda_v/\beta} - \left(\frac{\alpha_v \overline{\sigma}_v}{\omega_{vj,t+1} - C_{vj,t+1}}\right)^{\lambda_v/\beta}} > 0. \tag{1.21}$$

The derivative with respect to rainfall last period, $a_{v,t-1}$, is identical to $\partial \Delta \widehat{M}_{v,t+1}/\partial a_{vt}$ with a leading negative sign and shifting all t subscripts back to $t-1$. The proof that rainfall shocks have no effect in the absence of liquidity constraints is trivial since a_{vt} and $a_{v,t-1}$ do not enter equation (1.6). The positive effect of price shocks on $\widehat{M}_{v,t+1}$ in the presence of liquidity constraints follows immediately from the fact that $\lambda_v/\beta > 0$. The proof that price shocks have a negative effect on flow migration rates in the absence of liquidity constraints proceeds by checking that the following expression satisfies increasing differences (over time) in (H_{vs}, p_{vs}),

$$\ln\left[1 - (H_{vs} p_{vs})^{\frac{\lambda_v}{\beta}}\right],$$

where $H_{vs} = \alpha_v \overline{\sigma}_v K_v^{\theta} S_v^{\phi}/(\omega_{vj,t+1} - C_{vj,t+1})$. This condition holds so long as migration costs are non-increasing, $C_{vj,t+1} \leq C_{vjt}$. Of course, taking the derivative with respect to the price *level*, we find

$$\frac{\partial \Delta \widehat{M}_{v,t+1}}{\partial p_{vt}} = \frac{-\frac{\lambda_v}{\beta} p_{vt}^{\lambda_v/\beta - 1} \left(\frac{\alpha_v \overline{\sigma}_v K_v^{\theta} S_v^{\phi}}{\omega_{vj,t+1} - C_{vj,t+1}}\right)^{\lambda_v/\beta}}{1 - \left(\frac{\alpha_v p_{vt} \overline{\sigma}_v K_v^{\theta} S_v^{\phi}}{\omega_{vj,t+1} - C_{vj,t+1}}\right)^{\lambda_v/\beta}} < 0. \tag{1.22}$$

Referring instead to the equations in levels (1.3) and (1.5), the qualitative conclusions remain unchanged, i.e. the (monotone) comparative statics deliver the same predictions regarding the direction of the effect of rainfall and price shocks on flow migration rates.[57]

Proposition 2

The fact that λ_v has an ambiguous effect on the intensive margin follows immediately from differentiating equations (1.4) or (1.6) and recognizing that the terms inside brackets $[\cdot]$ within the logarithm are less than one. That $\frac{\partial \Delta \widehat{M}_{v,t+1}}{\partial \Delta \ln p_{vt} \partial \lambda_v} = 1/\beta > 0$ in the presence of liquidity constraints is immediate from equation (1.4). To show that $\partial^2 \Delta \widehat{M}_{v,t+1} / \partial a_{vt} \partial \lambda_v > 0$, simply rearrange and differentiate equation (1.21) with respect to λ_v

$$\begin{aligned} \frac{\partial^2 \Delta \widehat{M}_{v,t+1}}{\partial a_{vt} \partial \lambda_v} = & \frac{\frac{1}{\beta}(\overline{\sigma}_v + a_{vt})^{-1}}{1 - \left(\frac{\alpha_v \overline{\sigma}_v \tau_{vj} C_{vj,t+1}}{(\overline{\sigma}_v + a_{vt})(\omega_{vj,t+1} - C_{vj,t+1})}\right)^{\frac{\lambda_v}{\beta}}} \\ & + \frac{\frac{\lambda_v}{\beta^2}(\overline{\sigma}_v + a_{vt})^{-1} \left(\frac{\alpha_v \overline{\sigma}_v \tau_{vj} C_{vj,t+1}}{(\overline{\sigma}_v + a_{vt})(\omega_{vj,t+1} - C_{vj,t+1})}\right)^{\frac{\lambda_v}{\beta}} \ln \left(\frac{\alpha_v \overline{\sigma}_v \tau_{vj} C_{vj,t+1}}{(\overline{\sigma}_v + a_{vt})(\omega_{vj,t+1} - C_{vj,t+1})}\right)}{\left(1 - \left(\frac{\alpha_v \overline{\sigma}_v \tau_{vj} C_{vj,t+1}}{(\overline{\sigma}_v + a_{vt})(\omega_{vj,t+1} - C_{vj,t+1})}\right)^{\frac{\lambda_v}{\beta}}\right)^2}. \end{aligned} \tag{1.23}$$

Letting $x_v := \alpha_v \overline{\sigma}_v \tau_{vj} C_{vj,t+1}$ and $y_v := (\overline{\sigma}_v + a_{vt})(\omega_{vj,t+1} - C_{vj,t+1})$, recognizing that $x_v < y_v$ (for those migrating, i.e. $R_{iv} \in [R_L, R_U]$), and noting that $(y_v/x_v)^{\lambda_v/\beta} + (\lambda_v/\beta)\ln(x_v/y_v) > 1$, it can be shown that equation (1.23) is pos-

[57] The derivations are more complicated if prices (i) follow a higher-order autoregressive process or (ii) have a forecastable nonzero drift term, and/or (iii) households do not have rational expectations over the high frequency seasonality in prices. Nevertheless, the assumptions in Section 1.3.1 are largely consistent with the time series properties of rainfall and rice prices in Indonesia (and presumably elsewhere). Moreover, the first-order price formulation is sufficiently general to comprise more higher-order Markov processes (see Chambers and Bailey, 1996).

itive. In the absence of liquidity constraints, a similar calculation on equation (1.22) shows that $\partial^2 \Delta \widehat{M}_{v,t+1}/\partial p_{vt} \partial \lambda_v < 0$.

Extensive margin

Although not offered as a formal proposition, λ_v has an ambiguous effect on the extensive margin regardless of the formulation of the extreme landholding statistics. Under the finite sample formulation, the proof follows immediately from the derivative of equation (1.7) with respect to λ_v, $N_v R_U^{-\lambda_v N_v} \ln R_U - N_v(1 - R_L^{\lambda_v})^{N_v} R_L^{-\lambda_v} \ln R_L$, the sign of which cannot be determined without imposing *ad hoc* bounds on parameter values. The ambiguity similarly holds for the population-based order statistic approach in equation (1.8). Meanwhile, the positive effect of population size N_v on the extensive margin follows from straightforward differentiation.

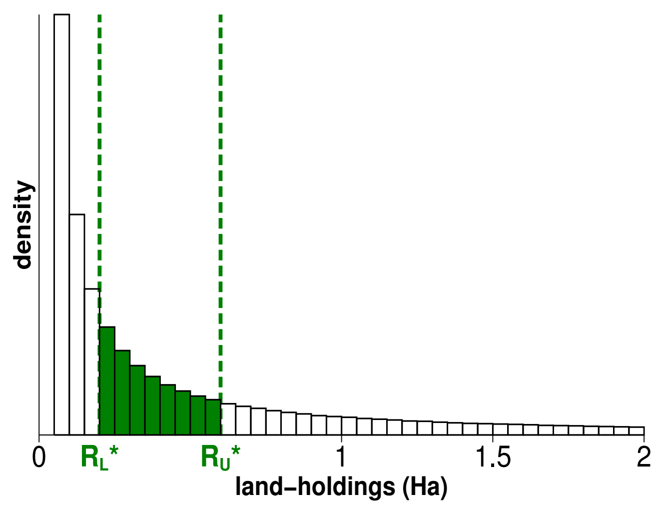


Figure 1.9: Stock Migration Rate for Given Triplet (R_L, R_U, λ_v)

1.9.2 Econometric Procedures

This appendix details two-step estimating framework introduced in Section 1.4.

Parametric

The parametric approach due to Poirier (1980) presumes that $(u_{vt}, u_{v,t+1}, \Delta\varepsilon_{v,t+1})$ in equation (1.10) follow a trivariate normal distribution with mean zero, variances $(1, 1, \text{var}(\Delta\varepsilon))$, and pairwise correlation terms $(\rho_{u_t u_{t+1}},\ \rho_{u_t \varepsilon},\ \rho_{u_{t+1}\Delta\varepsilon})$. These assumptions imply that

$$\mathbb{E}\left[\Delta\varepsilon_{v,t+1}\middle|\eta'_{t-1}\mathbf{Z}_{v,t-1} > -u_{vt}, \eta'_t\mathbf{Z}_{vt} > -u_{v,t+1}\right] = \rho_{u_t\Delta\varepsilon}\kappa_{vt} + \rho_{u_{t+1}\Delta\varepsilon}\kappa_{v,t+1},$$

where (suppressing v subscripts) the bivariate Mills ratio terms

$$\begin{aligned}
\kappa_t &= \phi(\eta'_{t-1}\mathbf{Z}_{t-1})\frac{\Phi\left(\frac{\eta'_t\mathbf{Z}_t - \rho_{u_t u_{t+1}}\eta'_{t-1}\mathbf{Z}_{t-1}}{\sqrt{1-\rho^2_{u_t u_{t+1}}}}\right)}{\Phi_2\left(\eta'_{t-1}\mathbf{Z}_{t-1}, \eta'_t\mathbf{Z}_t; \rho_{u_t u_{t+1}}\right)} \\
\kappa_{t+1} &= \phi(\eta'_t\mathbf{Z}_t)\frac{\Phi\left(\frac{\eta'_{t-1}\mathbf{Z}_{t-1} - \rho_{u_t u_{t+1}}\eta'_t\mathbf{Z}_t}{\sqrt{1-\rho^2_{u_t u_{t+1}}}}\right)}{\Phi_2\left(\eta'_{t-1}\mathbf{Z}_{t-1}, \eta'_t\mathbf{Z}_t; \rho_{u_t u_{t+1}}\right)}
\end{aligned}$$

with Φ and Φ_2 being, respectively, the univariate and bivariate standard normal CDF. Implementation proceeds in two steps. In the first step, I estimate a bivariate probit model for the extensive margins in t and $t+1$. In theory, identifiability of η_t and η_{t+1} merely requires that one element of $\mathbf{Z}_{t-1}$ and $\mathbf{Z}_t$ vary separately across villages in each equation (see Wilde, 2000). In practice, imposing at least one exclusion restriction across equations will dramatically improve the robustness of the identification. Since prices, rainfall and population size vary over time, the bivariate first stage is rich in sequential exclusion restrictions. Imposing theoretically motivated exclusion restrictions further the burden of identification on the mere nonlinearity of the κ terms. In the second step, I augment an empirical specification for the log flow migration rate with the estimated correction terms $\widehat{\kappa}_{vt}$ and $\widehat{\kappa}_{v,t+1}$, which enter with population coefficients equal to $\rho_{u_t\Delta\varepsilon}$ and $\rho_{u_{t+1}\Delta\varepsilon}$ respectively. Straightforward OLS then delivers a consistent estimate of second stage parameters. See Rochina-Barrachina (1999) for further theroetical background on the relationship between Poirier's original cross-sectional bivariate probit with partial observability and the two-period panel implementation as described here.

Semiparametric

Although tractable, the parametric selection model has strong distributional assumptions, departures from which can give rise to substantial specification biases. Following De Luca and Peracchi (forthcoming) and Das et al. (2003), this section outlines a practical semiparametric procedure for estimating the system of equations in (1.10) that is arguably more robust to distributional misspecification. Rather than closed-form correction terms, the semiparametric approach relies on a double-index in the propensity scores $g(\eta_{t-1}\mathbf{Z}_{v,t-1}, \eta_t\mathbf{Z}_{vt})$, where g is an unknown function of the latent variable indices.

Implementation proceeds as follows. First, rather than assuming bivariate normality of $(u_{vt}, u_{v,t+1})$, I (i) employ a semi-nonparametric pseudo-maximum likelihood (SNP-ML) procedure that uses an approximation to the unknown latent error densities (Gallant and Nychka, 1987), or (ii) estimate seemingly unrelated linear probability models (SU-LPM) making no assumptions on the joint distribution of u_{vt} and $u_{v,t+1}$ (Zellner and Lee, 1965). In both cases, there are no restrictions on the possible correlations between the error terms in t and $t+1$. The bivariate choice model in framework (i) is estimated according to a polynomial expansion where the orders of the expansion, L_t and L_{t+1}, for the two equations approximate the departures of u_t and u_{t+1} from normality.[58] Letting $\dim(\mathbf{Z}_{t-1}) = k_{t-1}$ and $\dim(\mathbf{Z}_t) = k_t$, identifiability of η_t and η_{t+1} requires that the vector $\mathbf{Z} = (\mathbf{Z}_{t-1}, \mathbf{Z}_t)$ assume at least $k_{t-1} + k_t + L_tL_{t+1}$ distinct values. This condition clearly holds given the large number of continuous variables (e.g., distance and population) determining the extensive margin of migration. Despite theoretical advantages, the SNP-ML estimates are qualitatively similar to those obtained with a SU-LPM pro-

[58] See De Luca and Peracchi (forthcoming) for an accessible introduction.

cedure. Give the vast differences in computational costs, I focus on the SU-LPM estimator beyond the baseline estimates in Tables 1.3-1.5.

Second, I use the estimates of η_t and η_{t+1} to approximate $g(\cdot)$. In practice, I employ an Lth-degree power series expansion in the propensity scores $\widehat{P}_s = \widehat{\eta}_s \mathbf{Z}_s$—linear predictions recovered from the bivariate SNP-ML or SU-LPM estimator—for village v to have at least one migrant in period s.[59] Lastly, consistent estimates of ζ can be obtained from the following OLS regression

$$\Delta \ln\left(\frac{M_{v,t+1}}{N_{v,t+1}}\right) = \boldsymbol{\zeta}' \Delta \mathbf{X}_{vt} + \sum_{p=0}^{L_t} \sum_{q=0}^{L_{t+1}} \mu_{pq} \left(\widehat{P}_{vt}\right)^p \left(\widehat{P}_{v,t+1}\right)^q + \Delta\varepsilon_{v,t+1} \tag{1.24}$$

since $\mathbb{E}[\Delta\varepsilon_{v,t+1}|\cdot] = 0$ so long as at least two variables in $\mathbf{Z}_{t-1} \cup \mathbf{Z}_t$ do not also appear in $\mathbf{X}_t$.

Inference

In both the parametric and semiparametric framework outlined above, the correction terms introduce added sampling variation into the second-stage.[60] Taking a conservative and unbiased approach to inference, I implement a bootstrap$-t$

[59] This is essentially the approach suggested by Das et al. (2003) who recommend using a fully nonparametric estimator to estimate the propensity score. Newey (1988) argues that a first stage linear probability model provides consistent estimates in two-step selection models, though a semiparametric first stage estimator provides more efficient (second-stage) estimates (Newey, 2009). An important difference with Das et al., however, is that they assume $u_{vt} \perp u_{v,t+1}$ whereas the estimates of η obtained using bivariate SNP-ML or SU-LPM explicitly allow for $\text{corr}(u_{vt}, u_{v,t+1}) \neq 0$. Results are robust to estimating two distinct LPMs with $\text{corr}(u_{vt}, u_{v,t+1}) = 0$.

[60] Although the Pareto parameters $\widehat{\lambda}_v$ are also estimated in auxiliary regressions, these fitted distributional terms are obtained from more than 55,000 regressions using the agricultural population of every Indonesian village in 2003 (more than 25 million households). There are a few reasons to be unconcerned with the added sampling variation introduced by these generated regressors. First, the Agricultural Census of 2003 purports to capture every agricultural household in rural Indonesia, and hence the distributional parameters are based on approximately population-level data. Second, these estimates are based on a different sample, at the individual-rather than the village-level as in the main regressions. Lastly, the consistency of the estimated λ_v holds under quite general assumptions (Gabaix and Ibragimov, 2011) and moreover should not affect the consistency of the estimates in either the first- or second-step of the two-step models (Newey and McFadden, 1994).

procedure (also known as percentile$-t$) with clustering at the district level.[61] All tables report the uncorrected standard errors, but the significance levels are computed based on the cluster bootstrap$-t$ procedure described in detail in Cameron et al. (2008a).[62] Each second-step significance level is based on 999 bootstrap iterations,[63] where I cluster the standard errors at each iteration and construct the iteration-specific Wald test statistic (t-stat) re-centered on the original point estimate. Using these 999 Wald statistics, I then compute the (possibly asymmetric) 90th, 95th, and 99th% confidence intervals in reporting the significance level $\alpha \in \{0.1, 0.05, 0.01\}$ associated with each point estimate.

The simulation results in Cameron et al. (2008a) suggest that the empirical setup in this paper is well suited to the cluster bootstrap$-t$ procedure. In particular, the data comprise a large number of districts (>200 in all specifications) with an unbalanced number of villages, several observable variables are relatively constant within district, and there are several binary regressors. Moreover, Yamagata (2006) finds that the bootstrap$-t$ procedure outperforms the conventional bootstrap$-se$ procedure in the context of estimating Heckman (1976)-type selection models similar to those in this paper.[64]

[61] Inference is robust to two-way clustering (see Cameron et al., 2011) on (i) the locations at which rainfall and price data are recorded, or (ii) district and plurality destination.

[62] In applications of the bootstrap$-t$ procedure, authors sometimes report p-values. While retaining the original biased standard errors, I report the unbiased significance levels where those p-values fall below 0.1. The underlying p-values are available upon request.

[63] The semiparametric correction procedure estimates based on SNP-ML in Table 1.5 do not have correct confidence intervals as bootstrapping is computationally impractical.

[64] The cluster bootstrap$-t$ procedure that I employ yields confidence intervals with correct coverage in addition to asymptotic refinement. In unreported results similar to Yamagata (2006), I also find that the 95% confidence intervals generated by a conventional cluster bootstrap$-se$ procedure fail to cover the original point estimate in more than 5% of iterations, suggesting important finite-sample shortcomings in line with other studies.

1.9.3 Data Description

Variable	Source	Definition
population	*Podes* 2005/8	all people registered as residents for at least six months or less than six months with the intention of staying
migrants	*Podes* 2005/8	all people working abroad on a fixed wage for a fixed time period
$\widehat{\lambda}_v$	Agricultural Census 2003	estimate of the Pareto exponent λ_v for village v based on OLS estimation (Gabaix and Ibragimov, 2011); see Appendix 1.9.6 for details
share households above $\underline{R}$	Agricultural Census 2003	share of all households in village v reporting landholdings less than $\underline{R}$ where $\underline{R} = 0.1$ hectares in the baseline case
rice prices	Wimanda (2009) via BPS	see Appendix 1.9.5 for details
rainfall in year t	NOAA/GPCP	total amount of rainfall during the given growing/harvest season where (1) seasons are 12 month intervals beginning with the first month of the province-specific wet season in a given year (Maccini and Yang, 2009), and (2) rainfall at the village level is based on rainfall levels recorded interpolated down to 0.5 degree (latitude/longitude) pixels between rainfall stations
net consumer/producer status, majority of HH	*Podes* 2005	whether head of village i reports in Podes 2005 that a majority of households in the village sell (net producers) or subsist (net consumers) on their agricultural production conditional on that village reporting agriculture being the most prominent source of employment
no motorized land travel to district capital	*Podes* 2002	equals one if there is no direct travel to the district capital using motorized land-based vehicles
reporting frequency	*Podes* 2005	assumes one of the following ordered levels: no formal population register, non-routine reporting, annual reporting, quarterly reporting, monthly reporting
distance to nearest district capital	*Podes* 2005/8	the minimum of the travel distance in kilometers to the given district capital or the nearest capital in a neighboring district
distance to subdistrict capital	*Podes* 2005/8	travel distance in kilometers to the capital of the village's subdistrict
distance to nearest emigration center		great circle distance from the centroid of the district in which village is located to the centroid of the nearest of 17 cities capable of processing legal international contract migration; cities include Aceh, Medan, Pekanbaru, Palembang, Jakarta, Bandung, Semarang, Yogyakarta, Surabaya, Pontianak, Banjarbaru, Nunukan, Makassar, Mataram, Kupang, Tanjung Pinang, and Bali
urban	*Podes* 2005/8	a government-constructed indicator which equals one if the village has a population density greater than 5000 per square kilometer, a majority of the population recorded as farming households and any number of public institutions which I do not observe directly in *Podes*
distance to Ho Chi Minh City/Bangkok (port)		great circle distance from the centroid of the village is located to the nearest Indonesian port plus the shipping distance abroad; geocoordinates of Indonesian port cities obtained from AtoBviaC and shipping distances obtained from e-ships
Arab (Chinese) population share	Population Census 2000	the number of individuals claiming Arab (Chinese) descent as a share of village population
Muslim population share	*Podes* 2005/8	the number of individuals claiming adherence to Islamic faith as share of village population
post-primary education share	Population Census 2000	share of the population aged 5 and above that has completed junior secondary (*SLTP/setara*), senior secondary (*SLTA/setara*), *or* post-secondary (*Diploma/DIII/Akedemi/DII/DIV*)
share population aged 15-29	Population Census 2000	age range is chosen to correspond to the majority migration age of 18-34 in later years as reported in the Bank Indonesia (2009) survey
estimated mean household expenditure/capita	Suryahadi et al. (2005)	estimate of the average household expenditures per month, obtained from the poverty mapping exercise based on the 2000 Census
total rice output in tons per Ha	*Podes* 2002	total rice output recorded in village in 2001 divided by total area harvested
bank presence	*Podes* 2002/5	all formal banking institutions including rural people's banks (BPR) and commercial microfinance (BRI)
village land area	*Podes* 2005/8	total land area in hectares

1.9.4 Further Empirical Results

This appendix presents several additional results and robustness checks on the main empirical findings in the paper.

Controlling for Negative Environmental Shocks

Several authors have demonstrated the impact of negative environmental shocks such as earthquakes in driving international migration (see, e.g., Halliday, 2006; Yang, 2008b). Natural disasters are commonplace in Indonesia. Utilizing data on the occurrence of mudslides, floods, earthquakes, fires, and other natural disasters recorded in *Podes* 2005 and 2008, I construct indicators for whether the village experienced more or less natural disasters in 2006-8 relative to 2003-5 with the reference category being no change in the number of given disasters.[65] Table 1.13 shows (for wetland holdings) that controlling for such disasters does not alter the key conclusions drawn in Sections 5.2 and 5.3. The baseline estimated impacts of rainfall and price shocks in column 5 are little affected by environmental shocks. The elasticities fall somewhat when controlling separately for all types of natural disasters, but the point estimates in columns 2-4 do not differ statistically from the original $\widehat{\theta}_p = 1.287$ and $\widehat{\theta}_a = 0.305$. Similarly, the estimates of $\theta_{a\lambda}$ and $\theta_{a\lambda}$ in columns 6-8 are relatively robust to controlling for intertemporal differences in the occurrence of any natural disasters.

More generally, the results in Table 1.13 indicate the strong effects of natural disasters on international migration flows. Villages experiencing more natural disasters in period t relative to $t-1$ tend to have higher migration flows. This might be explained by destruction of physical capital and farmland leading to a decline in the returns to labor at home. Another explanation could be that households use international migration as an *ex post* risk management strategy (as in Halliday, 2006). At the same time, villages experiencing fewer natural disasters in t than $t-1$ have lower migration outflows. This could be due in part to an influx

[65] Key results are unchanged when including an exhaustive set of indicators for the total difference rather than the binary more or less.

of outside capital and aid in the wake of natural disasters, often raising the returns to low-skilled labor used intensively in the rebuilding process. The precise causal pathway is beyond the scope of the theoretical model (and data) in the present study.

Controlling for Other Commodity Price and Regional Agricultural GDP Shocks

In this subsection, I consider how controlling for (i) other agricultural commodity price and (ii) overall regional agricultural GDP shocks affect the (interpretation of the) estimated relationship between rice price shocks, rainfall shocks, landholdings heterogeneity, and international migration flows. Although the rice price shock seen in Figures 1 and 7 in the paper was indeed large and invariably affected large swathes of the Indonesian population, several other agricultural products also exhibited some price volatility over the same period (crop-specific details are available upon request). To capture exposure at the village level to various price shocks, I build a commodity price index based on the prevailing mix of crops produced in a base year (2001 as reported in *Podes*) and *national average* farmgate prices for 44 of the most important cash and food crops other than rice grown in Indonesia from 2000-2008. Unlike the rice price shock, I am unable to exploit spatial variation in prices. Instead, I use variation in revealed crop choice prior to the sample period[66] to construct a non-rice agricultural commodity price index for village v in year t as a geometric average (see Bazzi and Blattman, 2012), $\mathcal{P}_{vt} = \prod_{k=1}^{44} \frac{p_{kt}^{s_{vko}}}{CPI_t}$, where p_{kt} is the national average farmgate price of crop k in year t available from the FAO, s_{vko} is the share of crop k in total agricultural GDP of

[66] I assume that all villages are price-takers in the domestic product market and that the crop-specific shares of village agricultural GDP are predetermined with respect to international migration in 2005/8. The former assumption can be tested heuristically. In terms of important food crops, no village produces more than one percent of total national output. The largest rice output (though not in the index) from a single village merely accounts for 0.03% of national output. The same does not hold as strongly for certain cash crops such as nutmeg/cardamom, cotton, pepper, tea, vanilla, and cinnamon for which a few villages across the country produce in excess of 5 percent of total national output. Excluding these village-crop observations from the price index does not affect any of the main empirical results in this section.

village v in base year o, and CPI_t is the Indonesian consumer price index in year t.[67]

In Table 1.14, I augment the baseline specifications in Tables 5 and 6 with the difference in log growth of the price index between t (year end 2002-2004) and $t+1$ (year end 2005-2007) (i.e., analogous to the rice price shock formulation). Controlling for other commodity price shocks besides rice does not affect the main conclusions drawn in the paper: $\widehat{\theta}_a$ and $\widehat{\theta}_{a\lambda}$ are positive and in most cases statistically significant, and $\widehat{\theta}_{p\lambda}$ is positive and statistically significant when restricting to $\widehat{\lambda}_v$ to wetland or paddy area planted. However, whereas $\widehat{\theta}_p$ is uniformly positive if not significant in Table 5, controlling for other price shocks in columns 2-3 of Table 1.14 renders the own rice price shock coefficient statistically insignificant throughout and sometimes flips its sign. The robust positive effect of other, non-rice commodity price shocks in row four suggests that the positive effect of rice price shocks on flow migration rates might be explained in part by the correlation between rice prices and other important agricultural commodities produced in rural Indonesia. The positive effect of those other price shocks on migration flows is nevertheless still consistent with the presence of liquidity constraints in the context of the theoretical model. Nevertheless, the estimated coefficients on the interaction terms, $\widehat{\theta}_{a\lambda}$ and $\widehat{\theta}_{p\lambda}$, remain similar to the baseline in Tables 5 and 6 with the slight exception that $\widehat{\theta}_{p\lambda}$ is relatively smaller when controlling for other commodity price shocks.

In Table 1.15, I demonstrate the robustness of the key results to controlling for a measure of overall regional agricultural GDP shocks. In particular, I augment the baseline specifications in Tables 5 and 6 with the difference in the log growth of agricultural GDP/capita at the district level between t and $t+1$.[68] Clearly,

---

[67] The weights are constructed as $s_{vko} = p_{k,2001} \times y_{kv,2001} / \left[\sum_{l=1}^{44} (p_{l,2001} \times y_{lv,2001})\right]$ where $p_{l,2001}$ is the average farmgate price of crop l in 2001 and $y_{lv,2001}$ is the output of crop l in the same year. A list of crops is provided in Appendix C in the paper.

[68] The agricultural GDP figures are reported in nominal *Rupiah* value by the Central Bureau of Statistics. Here, I use the logarithm of the values in 2002, 2005, and 2008. Some districts have missing data for these years and hence the slight difference in sample size with the baseline estimates in Tables 5 and 6 in the paper.

agricultural GDP should be highly collinear with rainfall as well as the general vitality of the rice economy. Thus, it is not surprising to find that our estimates of the key Θ parameters are slightly altered by accounting for regional shocks to overall agricultural productivity. Not unlike the previous Table 1.14, the estimates of θ_p are attenuated when controlling for a general measure of agricultural GDP shocks (regardless of the measure of $\widehat{\lambda}_v$ across the three panels). However, the key estimates on the interaction terms, $\widehat{\theta}_{a\lambda}$ and $\widehat{\theta}_{p\lambda}$, remain qualitatively similar to albeit slightly smaller in magnitude than the baseline estimates in Tables 5 and 6. Analogous to Table 1.14, we also find that flow migration rates are increasing in positive agricultural GDP shocks—a finding which is consistent with binding liquidity constraints given the intuition underlying the theoretical model in the paper.[69] However, unlike the rainfall and rice price shocks, it is difficult to know whether the residual agricultural GDP shock is comprised of variation that is endogenous with respect to international migration flows. Rainfall shocks are by definition exogenous, and I took steps to control for the potentially endogenous component of the price shock.

Alternative Specifications of Rainfall and Rice Price Shocks

Table 1.16 shows that the primary conclusions regarding the effects of rainfall shocks are robust to the inclusion of period-specific shocks rather than the difference in shocks between t and $t-1$. Furthermore, I fail to reject that the coefficient on the rainfall shock in t equals the absolute value of the coefficient on the rainfall shock in $t-1$.

[69] I find similar results for Θ and the own coefficient on a measure of non-agricultural (non-oil) GDP shocks, which is unsurprising given that the measure is highly correlated (above 0.9) with agricultural GDP shocks. The results are similar when controlling for both agricultural and non-agricultural GDP shocks simultaneously. However, to the extent that such estimates are not confounded by multicollinearity, I find that agricultural GDP shocks increase migration flows whereas non-agricultural GDP shocks decrease migration flows. One plausible explanation for this finding would be that non-agricultural GDP shocks are highly correlated with labor demand shocks, and those villages with some (light) industrial activity experience an influx of internal migrants from other villages thereby increasing the denominator (village population size) in the dependent variable—akin to the effect of rainfall shocks on internal mobility (see Section 1.9.4 below). These results are available upon request.

Table 1.17 considers an alternative specification for rainfall shocks in which the annual shocks are fully elaborated from 2002-8 (i.e., the rainfall shock in each season s is assigned its own elasticity parameter θ_{as} for $200s = 3, \ldots, 8$ and $\theta_{as\lambda}$ for the interactions with $\widehat{\lambda}_v$). At the bottom of the table, I report the sum of coefficients for period t ($s = 3, 4, 5$), period $t-1$ ($s = 6, 7, 8$), and both t and $t-1$ ($s = 3, \ldots, 8$) as well as the associated p-value for the null hypothesis that the given sum equals zero. In columns 1-3, we draw the same conclusions as in Table 1.16: (i) the sum of period t ($t-1$) rainfall shocks is positive (negative) and statistically significant, and (ii) the null hypothesis that $\theta_{a3} + \theta_{a4} + \theta_{a5} = -(\theta_{a6} + \theta_{a7} + \theta_{a8})$ cannot be rejected. Furthermore, in columns 4-6, we similarly rule out the possibility that the baseline specification of rainfall shocks leads to spurious conclusions regarding the key elasticity parameter $\theta_{a\lambda}$. That is, the sum of period t ($t-1$) coefficients on the interaction of rainfall shocks and $\widehat{\lambda}_v$ are positive (negative) and statistically significant.[70] In unreported results, I also show that the main results are robust to allowing negative rainfall shocks to have a different effect than positive rainfall shocks (i.e., rather than using a single continuous measure crossing zero).

Table 1.18 considers alternative approaches to measuring the rice price shock. Columns 1-4 report estimates of θ_p and $\theta_{p\lambda}$ using $\widehat{\lambda}_v$ for wetland holdings.[71] In columns 5-8, I specify the price "shock" as a difference in log *average* prices over 2005m4-2008m3 and 2002m2-2005m3 rather than a difference in annualized log growth rates between those two periods. This specification yields similar results. In columns 9-12, I adopt insights from the model for rice prices developed in Appendix 1.9.5. Because the model predicts that the price shock should be *decreasing* in distance from port cities in Indonesia and the shipping routes to

[70] An interesting feature of the fully elaborated specification is that the s and $s-1$ coefficients alternative in sign, with the period s contemporaneous with the *Podes* enumeration dates in 2005 and 2008 being positive. Two factors might explain this pattern: (i) the mean reverting properties of rainfall (see Appendix 1.9.5), and/or (ii) a particular spatial distribution of two-year migration contract cycles. Nevertheless, the cumulative migration flows are what we observe in the data and hence the sum is what matters, not the individual years per se.

[71] In Appendix 1.9.5, I argue against the use of another measure of rice prices available from unit values reported in household survey data.

Thailand and Vietnam, a negative coefficient on the two distance terms would be consistent with a positive elasticity of migration flows with respect to rice price shocks. Columns 9-10 are consistent with this hypothesized relationship as are the negative coefficients on the interaction terms with $\widehat{\lambda}_v$ in columns 11-12. Because all specifications control for log distance to the nearest legal emigration center, most of the identification in these columns comes from villages for which the nearest migration hub differs from the nearest port city.

Alternative Choices of $\underline{R}$

Although the λ_v parameters should be unaffected by the location of $\underline{R}$, in practice, the Pareto distribution is only an approximation, which works better in some villages than others (see Appendix 1.9.6). Nevertheless, in Table 1.19, I show that the estimates of the key elasticity parameters generally do not change when imposing alternative $\underline{R} \in \{0.15, 0.2, 0.25\}$ in the estimation of λ_v (and the share of households above $\underline{R}$) for wetland holdings. The results are similar for λ_v estimated using total agricultural landholdings or paddy area planted in 2002.[72]

Parsimony and Parallel Trends

One concern with using rainfall or price shocks is that they may not satisfy a parallel trend assumption. That is, the effects we observe may have been predetermined by other covariates in the model. If the shocks are truly exogenous, then their effect on migration should not hinge on which covariates are included. Table 1.20 employs the usual semiparametric and parametric correction procedures to demonstrate that the qualitative and quantitative results for both rainfall and price shocks are indeed insensitive to the choice of time-invariant covariates in the model.[73] Interestingly, in the most parsimonious specifications in columns 13-18, I find that the estimated elasticity of migration flows with respect to rainfall shocks

[72]In unreported results, I also show that similar conclusions hold in moving $\underline{R}$ up to 0.5 hectares. Note that the sample sizes differ across columns because consistent (i.e., usable) estimates of λ_v require at least 3 distinct size measures above $\underline{R}^*$. Some villages do not satisfy this criteria for a given minimum threshold value and landholding type.

[73]Note that this is by definition a partial test of the parallel trend identifying assumption (see Yang, 2006).

is unaffected by the exclusion of the rice price shock term whereas the elasticity with respect to rice price shocks doubles when not controlling for rainfall shocks (column 17). This suggests the importance of controlling simultaneously for correlated agricultural income shocks. Table 1.21 makes the same set of arguments for the estimates of $\theta_{a\lambda}$ and $\theta_{p\lambda}$ in Tables 7 and 8.

Accounting for Village Demographic Structure and Past Internal Migration

Table 1.22 demonstrates that the main results in Sections 5.2 and 5.3 in the paper are robust to controlling for (i) the share of the population that lived outside the village in 1995, (ii) the share of population aged 15-29, and (iii) the average household size in the village—each drawn from the 2000 Population Census. Variable (i) proxies for potential prior experience in and network connections to domestic labor markets outside the village. Variable (ii) captures to some extent labor market pressures induced by Indonesia's relatively recent demographic transition. Moreover, individuals within that given cohort are the most likely to have been potential migrants beginning 3-7 years later and hence recorded in 2005 and 2008 migrant stocks.[74] Although highly correlated with mean village income, mean household size also picks up variation in household labor supply, which may in turn affect the robustness of agricultural labor markets (i.e., off own-farm) and the capacity of households to diversify labor allocation across borders—both of which could have direct effects on flow migration rates.

On the (Non-)Effect of Measurement and Reporting Outliers

Tables 1.23 and 1.24 demonstrate that the key estimates of $\widetilde{\Theta} \equiv (\theta_a, \theta_p, \theta_{a\lambda}, \theta_{p\lambda})$ in the paper are robust to and arguably strengthened by accounting for outliers in the data along a few important dimensions. Column 2

[74] Of course, inclusion of this variable might introduce a source of bias in that villages with a large share of aged 15-29 in 2000 may be precisely those villages for which (i) the Asian financial crisis of 1997-8 led to a large return migration from urban areas, and/or (ii) the local economy was (expected to be) thriving as global agricultural commodity prices remained high through the early 2000s. Insomuch as these effects are persistent, the demographic variable could be endogenous and hence pose problems for the estimation of key parameters in this model.

controls for the frequency with which the village updates its population register (see Appendix C in the paper). This helps account for some of the measurement error in migration rates as well as potential misclassification bias arising from villages reporting no migrants when in fact there is at least one migrant from the village. Column 3 trims the bottom and top 1 percent of $\widehat{\lambda}_v$. Column 4 removes villages subject to censoring in reported migrant stocks in 2005 and/or 2008.[75] In column 5, I retain only those villages for which I did not have to rely on any fuzzy matching algorithms for merging villages across the 2005 and 2008 waves of *Podes* (see Section 1.9.8). Although I have confidence in the matching algorithms, they are by design imperfect and hence may contribute to measurement error on both sides of the estimating equation. Last, column 6 simultaneously implements the prior four restrictions. In all cases, the main qualitative and quantitative interpretation of Θ remains unchanged.

In column 7 of Tables 1.23 and 1.24, I drop provinces identified in Bank Indonesia (2009) as having a large number of undocumented international migrants (primarily going to Malaysia). The Village Potential data, recall from Section 2.1, merely define international labor migrants as those working abroad for a fixed wage and time period. It is possible therefore that this count includes some undocumented migrants for which the determinants of migration choice and the nature of liquidity constraints may be somewhat different than for legal migrants. When dropping these provinces—which, keep in mind, still have a large number of legal international migrants—a few differences emerge with respect to the full sample results. First, in Table 1.23, the elasticity parameters for rainfall and price shocks slightly increase. However, the estimates of $\lambda_{a\lambda}$ and $\lambda_{p\lambda}$ in column 7 fall in magnitude. The large, precisely estimated $\lambda_{p\lambda}$ for $\widehat{\lambda}_v$ based wetland holdings disappears entirely. It seems, then, that undocumented migrants may explain some of the

[75]The 2005 survey records separately the total number of male migrants and the total number of female migrants whereas the 2008 survey simply records the total number of migrants. Whether the different format of the question across years biases reporting is an open question. However, top coding poses a challenge in the following sense. In 2005, the separate reporting for male and female migrants allowed total migrant stocks to exceed 998 persons for 40 villages while villages could only record a maximum of 998 persons abroad in the 2008 survey.

stronger response of migration flows to price shocks in villages with a larger mass of small landholders.

Rainfall Shocks and Internal Migration

Here, I briefly discuss the effect of rainfall shocks on internal migration flows as mentioned in Section 5.5.1 in the paper. Using weighted samples from Indonesian Population Censuses in 2000 and 2010 as well as Intercensal Population Surveys in 1985, 1995, and 2005, I am able to construct a bilateral district-level migration matrix in which each observation comprises the stock of individuals hailing from origin district o in year $t-5$ and currently residing in destination district d in year t.[76] I estimate the following quasi-gravity model for internal (h for home) migration flows as a function of origin and destination rainfall shocks:

$$\ln migrants^h_{odt} = \alpha rainfall\ shock_{ot} + \beta rainfall\ shock_{dt} + \upsilon_o + \upsilon_d + \upsilon_t + \epsilon_{odt}. \quad (1.25)$$

where, for $j = o, d$, $rainfall\ shock_{jt}$ captures (in logarithmic form) the cumulative annual rainfall shocks over the four year prior to t,[77] υ_j geographic fixed effects, υ_t is a year fixed effect, ϵ_{odt} is an idiosyncratic error term.[78]

Estimating equation (1.25) by OLS for the entire period 1985-2010, I find $\widehat{\alpha} \approx -0.056$ (std. error of 0.022), which suggests that origin rainfall shocks reduce internal out-migration. Restricting to the period 2005-2010—roughly, the period over which I observe international migrants in the Village Potential data used in the paper—I obtain $\widehat{\alpha} \approx -0.452$ (std. error of 0.071).[79] (In both cases, I also find

[76]The data were downloaded from the Integrated Public Use Microdata Series, International in August 2012. The district-level total migrant and population counts are based on summing the person-specific population weights provided by IPUMS-International and representative at the district-level. Details on the (Inter-)Census specific samples can be obtained on the IPUMS website for Indonesia. Further details on the panel construction are available upon request.

[77]For example, the shock in for origin district k in 2005 is simply the sum of the annual log deviations in 2001-2004 from the long-run district-level mean calculated over all years from 1948-2010 excluding 2001-2004.

[78]We use the log number of migrants rather than the migrant share or the odds of migration quite simply because the goal is to characterize changes in district population levels arising from internal migration (i.e., the denominator in the dependent variable in the paper).

[79]I cluster standard errors by origin×destination district pair. Standard errors increase slightly

that $\beta > 0$ and statistically significant, which is consistent with migration being responsive to destination wage shocks.) Taken together, the negative estimates of α support the claim that positive rainfall shocks increase district population size and hence are likely also to increase village population size, presuming (i) inter-district migration is a lower bound for overall internal out-migration observed at the village level, and (ii) intra-district migration outside the home village follows similar processes. Such upward pressure on village population size in the denominator of the dependent variable in the paper (Δlog migrants/population) implies that the positive relationship between changes in *international* migration rates and rainfall cannot be explained by the unobservable internal migration flows at the village level.[80]

Further Background on the Validation Exercise Using Micro Data

In Section 5.6 of the paper, I discuss results from estimating a migration choice model and using the implied marginal effects to recover an alternative measure of the village-level elasticity of flow migration rates with respect to income shocks. In this brief subsection, I provide a few additional details on the analysis therein.

First, note that in columns 3-4 of Table 10 in the paper, I report coefficient estimates from the following equation

$$\begin{aligned} migrate_{iv,t+1} \quad = \quad & \alpha + \beta \; rainfall \; shock_{vt} + \gamma \; price \; shock_{vt} \\ & + price \; shock_{vt} \times (land_i \zeta_1^p + land_i^2 \zeta_2^p) \\ & + rainfall \; shock_{vt} \times (land_i \zeta_1^a + land_i^2 \zeta_2^a) \\ & + \eta_i + \; \eta_t + \; e_{iv,t+1}, \end{aligned}$$

which, recall, I estimate using a conditional fixed effects (CFE) logit estimator,

when using two-way clustering (Cameron et al., 2011) on both origin and destination district.

[80] In unreported results, I incorporate rice price shocks into equation (1.25). However, the data are only available for a subset of the entire period (2002-8) and hence are not as well-suited to the migration time horizons as are the rainfall data. When including them nevertheless, I find null results for the coefficient on origin price shocks and negative albeit imprecise estimates on destination price shocks.

and where (i) $land_i$ comprises all landholdings owned, under rental, or rented out and used to grow rice, and (ii) column 3 imposes $\zeta_2^a = 0$ and $\zeta_2^p = 0$. Using these estimates, I then recover average marginal effects (AMEs) at each value of $land_i \in \{0.1, 0.2, \ldots, 2.5\}$ Ha, where (i) 2.5 Ha is the maximum in the sample, and (ii) and the calculation of AMEs requires imposing $\eta_i = 0\ \forall i$. Thus, we obtain AMEs for both rainfall and rice price shocks at each landholding size (at 0.1 Ha increments).

As described in Section 5.6, I use these individual-level AMEs to obtain aggregate village-level elasticities of migration rates with respect to income shocks. I do so by applying the population shares to each landholding size-specific AME as implied by the village-level Pareto distribution. Consider, for example, the AMEs for rainfall shocks at landholding sizes 0.3 and 0.4 Ha. For each village v, I reweight the average of these two AMEs by the share of the population with landholding sizes$\in [0.3, 0.4]$ Ha as implied by the Pareto exponent $\widehat{\lambda}_v$.[81] I repeat this over all increments of landholding sizes in the village, apply the AME at 2.5 Ha to all households above 2.5 Ha (as implied by $\widehat{\lambda}_v$), and then sum the reweighted AMEs to recover an aggregate village-level elasticity. In Table 11, I then compared these implied elasticities to those from the actual village-level regressions, which allowed the effect of income shocks on flow migration rates to vary with $\widehat{\lambda}_v$.

In recovering the elasticities of flow migration rates with respect to price and rainfall shocks, I take the baseline coefficient estimates of $\widetilde{\Theta} \equiv (\theta_a, \theta_p, \theta_{a\lambda}, \theta_{p\lambda})$ in column 14 of Table 6 for $\widehat{\lambda}_v$ based on wetland holdings. I then assign to village v the average marginal effects of the price shock for all villages with $\widehat{\lambda}_v$ in the same percentile. That is, I calculate the average marginal effects of income shocks at each percentile of the distribution of $\widehat{\lambda}_v$ in the second-step sample of villages. Following this procedure makes it possible to compare the village-level elasticities

[81] One could also imagine reweighting nonparametrically by applying the observed shares in the Agricultural Census. The approach based on $\widehat{\lambda}_v$ is more consistent with the testable implications of the theoretical model and is moreover necessary for the purposes of comparison with the village-level elasticities of income shocks that vary with $\widehat{\lambda}_v$.

with analogous elasticities recovered from an underlying migration choice model.

Further Background on the Estimation of Village-Specific Migration Costs

Having found strong empirical evidence of financial constraints to migration in the empirical analysis of Sections 5.1-5.4, I used the following structural equation (4) for the log flow migration rate to back out estimates of the migration costs in Section 5.7 of the paper:

$$\Delta \ln \left(\frac{M_{v,t+1}}{N_{v,t+1}} \right) = \frac{\lambda_v}{\beta} \Delta \ln p_{vt} + \Delta \ln \left[\left(\frac{\overline{\sigma}_v + a_{vt}}{\tau_{vj} C_{vj,t+1}} \right)^{\frac{\lambda_v}{\beta}} - \left(\frac{\overline{\sigma}_v \alpha_v}{\omega_{vj,t+1} - C_{vj,t+1}} \right)^{\frac{\lambda_v}{\beta}} \right].$$

Here, I provide a few additional details on the calculation of these village-specific migration costs not mentioned in the paper.

First, I plug in the empirical analogues for rice prices and rainfall. I specify $\Delta \ln p_{vt}$ in the above equation as the log difference in the local rice price index over the entire period, 2002m1-2008m3. I set the rainfall level, $\sigma_{vt} \equiv \overline{\sigma}_v + a_{vt}$, equal to the average of the annual seasonal rainfall levels (in centimeters) over the three seasons prior to mid-2008 (mid-2005 for $\sigma_{v,t-1}$). I set $\overline{\sigma}_v$ equal to the average annual seasonal rainfall levels (in centimeters) over the 55 year period 1953-2008. The rainfall shocks a_{vt} capture the empirical difference between σ_{vt} and $\overline{\sigma}_v$.

Second, the prevailing wage offers ω_{vjt} described in the paper are, in fact, not village-specific on account of the facts that (i) the only information on destinations that I have at the village-level is the plurality choice of migrants from that village in mid-2005, and (ii) Bank Indonesia (2009) and other available sources merely report the monthly wages for low-skill Indonesian workers in each of the destination countries as stipulated in bilateral Memoranda of Understanding and reported by recruiters. These typical wages fall between the very narrow range of actual wages received as reported by migrants in the Bank Indonesia (2009) survey. Wages increased in early 2007 for most of the plurality destinations in the

Village Potential 2005 data, and for those that do not, I nevertheless increase the wages by 10 percent. The results are robust to other choices.

Plugging in the relevant empirical data into the above equation, I then solve for the fixed migration costs C_{vjt}. Obtaining an analytic solution, however, requires a few additional simplifications. First, I assume that migration costs are constant across periods. This assumption is conservative insomuch as migration costs likely fell in response to (i) competitive pressures in the recruitment industry and (ii) improvements in transportation infrastructure including the addition of new legal emigrant processing centers in a few provinces. Second, I assume that $\tau_{vj} = \overline{\tau}$ for all villages. However, because the estimates of C_{vjt} are sensitive to the choice of $\overline{\tau}$, I take the average C_{vjt} obtained for $\overline{\tau} \in \{0.01, 0.02, \ldots, 1\}$. Lastly, I impose $\omega_{vjt} = \omega_{vj,t+1} = \widehat{\omega}_{vj}$, and I set $\widehat{\omega}_{vj}$ to be the average of the empirical wages across both periods.

Tables

Table 1.13: Controlling for Other Negative Environmental Shocks

Landholdings measure	**Wetland Holdings**							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Δ rainfall shock	0.305	0.280	0.283	0.238	0.262	0.248	0.239	0.209
	(0.134)*	(0.135)*	(0.133)*	(0.133)*	(0.168)	(0.168)	(0.167)	(0.165)
Δ price shock	1.287	1.214	1.093	0.936	-2.234	-1.953	-2.356	-1.729
	(0.502)*	(0.515)*	(0.502)*	(0.489)*	(0.709)*	(0.695)	(0.715)**	(0.685)
Δ rainfall shock $\times$ $\widehat{\boldsymbol{\lambda}}_v$					0.028	0.023	0.031	0.026
					(0.052)	(0.050)	(0.050)	(0.048)
Δ price shock $\times$ $\widehat{\boldsymbol{\lambda}}_v$					1.913	1.727	1.871	1.456
					(0.335)***	(0.324)***	(0.334)***	(0.321)***
$\mathbf{1}(\Delta\ \text{earthquake}_{v,t+1} > 0)$		0.325				0.308		
		(0.051)***				(0.050)***		
$\mathbf{1}(\Delta\ \text{earthquake}_{v,t+1} < 0)$		0.019				0.028		
		(0.052)				(0.052)		
$\mathbf{1}(\Delta\ \text{floods or mudslides}_{v,t+1} > 0)$			0.202				0.191	
			(0.023)***				(0.023)***	
$\mathbf{1}(\Delta\ \text{floods or mudslides}_{v,t+1} < 0)$			-0.268				-0.258	
			(0.030)***				(0.030)***	
$\mathbf{1}(\Delta\ \text{total natural disasters}_{v,t+1} > 0)$				0.162				0.154
				(0.021)***				(0.021)***
$\mathbf{1}(\Delta\ \text{total natural disasters}_{v,t+1} < 0)$				-0.181				-0.176
				(0.025)***				(0.025)***
Number of villages	24,537	24,537	24,537	24,537	24,537	24,537	24,537	24,537

Landholdings measure	**Total Agricultural Landholdings**							
	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Δ rainfall shock	0.415	0.393	0.417	0.353	0.225	0.223	0.218	0.193
	(0.133)***	(0.134)**	(0.133)***	(0.132)**	(0.169)	(0.169)	(0.170)	(0.167)
Δ price shock	0.409	0.418	0.315	0.344	-0.016	0.086	-0.198	0.055
	(0.448)	(0.462)	(0.449)	(0.442)	(0.688)	(0.698)	(0.688)	(0.682)
Δ rainfall shock $\times$ $\widehat{\boldsymbol{\lambda}}_v$					0.119	0.105	0.125	0.101
					(0.074)*	(0.074)	(0.074)**	(0.073)*
Δ price shock $\times$ $\widehat{\boldsymbol{\lambda}}_v$					0.267	0.206	0.324	0.182
					(0.329)	(0.330)	(0.329)	(0.328)
$\mathbf{1}(\Delta\ \text{earthquake}_{v,t+1} > 0)$		0.348				0.341		
		(0.047)***				(0.047)***		
$\mathbf{1}(\Delta\ \text{earthquake}_{v,t+1} < 0)$		-0.061				-0.057		
		(0.062)				(0.062)		
$\mathbf{1}(\Delta\ \text{floods or mudslides}_{v,t+1} > 0)$			0.204				0.201	
			(0.024)***				(0.024)***	
$\mathbf{1}(\Delta\ \text{floods or mudslides}_{v,t+1} < 0)$			-0.234				-0.233	
			(0.027)***				(0.027)***	
$\mathbf{1}(\Delta\ \text{total natural disasters}_{v,t+1} > 0)$				0.166				0.163
				(0.021)***				(0.021)***
$\mathbf{1}(\Delta\ \text{total natural disasters}_{v,t+1} < 0)$				-0.177				-0.176
				(0.023)***				(0.023)***
Number of villages	26,527	26,527	26,527	26,527	26,527	26,527	26,527	26,527

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The table examines the robustness of key parameter estimates to the inclusion of controls for negative environmental shocks. The dependent variable in all specifications is $\Delta \ln(M_{v,t+1}/N_{v,t+1})$ and has mean 0.11. Sample sizes are identical across sub-columns within the super-column, as reported at the bottom of the table. Standard errors are clustered at the district level and significance levels are based on a block bootstrap$-t$ procedure. Total natural disasters includes forest fires and typhoons as well as earthquakes, floods, and mudslides. Additional covariates in all specifications but not reported here include all those under Table 5 in the paper.

Table 1.14: Controlling for Other Agricultural Commodity Price Shocks

Correction Procedure	None	Semipar.	Param.	None	Semipar.	Param.	None	Semipar.	Param.
1st Stage Estimator	None	SU-LPM	BiProbit	None	SU-LPM	BiProbit	None	SU-LPM	BiProbit
Landholdings measure					**Total Agricultural Landholdings**				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pareto exponent $\widehat{\lambda}_v$	-0.002	0.031	0.030	-0.002	0.013	0.028	0.045	0.011	0.009
	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.039)	(0.036)	(0.035)
Δ rainfall shock	0.100	0.364	0.258	0.071	0.148	0.104	0.107	0.195	0.164
	(0.127)	(0.133)**	(0.131)**	(0.166)	(0.169)	(0.161)	(0.166)	(0.171)	(0.164)
Δ rice price shock	-0.191	-0.071	-0.135	-0.191	-0.052	-0.131	0.572	-0.088	-0.442
	(0.445)	(0.440)	(0.435)	(0.446)	(0.438)	(0.436)	(0.750)	(0.697)	(0.701)
Δ other price shock	0.084	0.083	0.076	0.085	0.085	0.077	0.084	0.085	0.077
	(0.033)**	(0.034)	(0.033)*	(0.033)**	(0.034)*	(0.033)*	(0.033)**	(0.034)	(0.033)*
Δ rainfall shock $\times$ $\widehat{\lambda}_v$				0.019	0.138	0.102	-0.006	0.106	0.063
				(0.071)	(0.076)**	(0.070)*	(0.069)	(0.078)	(0.072)
Δ rice price shock $\times$ $\widehat{\lambda}_v$							-0.493	0.018	0.206
							(0.361)	(0.339)	(0.337)
Number of Villages	26,092	26,091	26,091	26,092	26,091	26,091	26,092	26,091	26,091
Landholdings measure					**Wetland Holdings**				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pareto exponent $\widehat{\lambda}_v$	-0.018	0.035	0.027	-0.019	0.036	0.028	0.005	-0.080	-0.054
	(0.014)	(0.017)*	(0.017)*	(0.015)	(0.018)**	(0.017)*	(0.022)	(0.029)**	(0.026)**
Δ rainfall shock	0.051	0.210	0.161	0.070	0.140	0.069	0.062	0.196	0.121
	(0.132)	(0.136)	(0.133)	(0.164)	(0.166)	(0.166)	(0.163)	(0.168)	(0.161)
Δ rice price shock	-0.211	0.234	-0.043	-0.212	0.239	-0.042	0.273	-1.926	-1.442
	(0.448)	(0.460)	(0.447)	(0.448)	(0.459)	(0.447)	(0.608)	(0.697)*	(0.660)**
Δ other price shock	0.082	0.116	0.100	0.082	0.116	0.099	0.082	0.115	0.098
	(0.033)**	(0.036)**	(0.034)***	(0.033)**	(0.036)**	(0.034)***	(0.033)**	(0.036)**	(0.034)**
Δ rainfall shock $\times$ $\widehat{\lambda}_v$				-0.011	0.041	0.055	-0.010	0.013	0.027
				(0.049)	(0.049)	(0.056)	(0.047)	(0.051)	(0.050)
Δ rice price shock $\times$ $\widehat{\lambda}_v$							-0.266	1.187	0.772
							(0.232)	(0.320)***	(0.292)***
Observations	24,540	24,537	24,537	24,540	24,537	24,537	24,540	24,537	24,537
Landholdings measure					**Paddy Area Planted, 2002-3**				
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pareto exponent $\widehat{\lambda}_v$	0.015	0.030	0.036	0.015	0.017	0.033	0.027	-0.060	-0.062
	(0.016)	(0.017)	(0.017)*	(0.016)	(0.017)	(0.017)*	(0.032)	(0.050)	(0.041)**
Δ rainfall shock	0.088	0.298	0.201	-0.095	0.052	-0.013	-0.089	0.101	0.050
	(0.133)	(0.142)*	(0.139)	(0.166)	(0.173)	(0.172)	(0.168)	(0.176)	(0.175)
Δ rice price shock	-0.230	0.122	-0.143	-0.214	0.126	-0.126	-0.000	-1.218	-1.751
	(0.438)	(0.455)	(0.445)	(0.440)	(0.455)	(0.445)	(0.699)	(0.853)	(0.774)**
Δ other price shock	0.093	0.096	0.090	0.093	0.095	0.091	0.093	0.094	0.091
	(0.033)***	(0.035)**	(0.033)**	(0.033)***	(0.034)**	(0.033)***	(0.033)***	(0.034)*	(0.033)**
Δ rainfall shock $\times$ $\widehat{\lambda}_v$				0.113	0.145	0.130	0.108	0.123	0.095
				(0.053)**	(0.060)**	(0.057)**	(0.054)**	(0.062)*	(0.059)*
Δ rice price shock $\times$ $\widehat{\lambda}_v$							-0.134	0.766	0.985
							(0.295)	(0.452)*	(0.385)***
Observations	24,476	24,476	24,476	24,476	24,476	24,476	24,476	24,476	24,476

Notes: The table reports estimates analogous to those in Tables 5 and 6 in the paper but augmented with a variable capturing other agricultural commodity price shocks as described in Section 1.9.4. The dependent variable in all specifications is $\Delta \ln M_{v,t+1}/N_{v,t+1}$ and has mean 0.11. Standard errors are clustered at the district level and significance levels are based on a block bootstrap$-t$ procedure. Additional covariates in all specifications but not reported here include all those reported under Table 5 in the paper. See Appendix B for details on the semiparametric (Semipar.) and parametric (Param.) correction procedures.

Table 1.15: Controlling for Overall Agricultural GDP Shocks

Correction Procedure	None	Semipar.	Param.	None	Semipar.	Param.	None	Semipar.	Param.
1st Stage Estimator	None	SU-LPM	BiProbit	None	SU-LPM	BiProbit	None	SU-LPM	BiProbit
Landholdings measure				**Total Agricultural Landholdings**					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pareto exponent $\widehat{\lambda}_v$	-0.001	0.032	0.031	0.000	0.013	0.030	0.045	0.005	0.009
	(0.018)	(0.018)	(0.018)	(0.019)	(0.019)	(0.019)	(0.039)	(0.036)	(0.037)
Δ rainfall shock	0.137	0.373	0.279	0.102	0.174	0.130	0.136	0.237	0.199
	(0.138)	(0.144)**	(0.135)**	(0.184)	(0.188)	(0.179)	(0.183)	(0.189)	(0.181)
Δ rice price shock	-0.217	-0.127	-0.155	-0.217	-0.118	-0.156	0.515	-0.222	-0.476
	(0.437)	(0.437)	(0.432)	(0.438)	(0.436)	(0.434)	(0.745)	(0.696)	(0.722)
Δ agri. GDP/capita shock	0.049	1.082	0.766	0.048	1.052	0.760	0.048	1.042	0.758
	(0.148)	(0.174)***	(0.156)***	(0.148)	(0.175)***	(0.156)***	(0.148)	(0.175)***	(0.156)***
Δ rainfall shock $\times\ \widehat{\lambda}_v$				0.023	0.126	0.099	-0.002	0.083	0.053
				(0.073)	(0.076)**	(0.072)*	(0.071)	(0.075)	(0.073)
Δ rice price shock $\times\ \widehat{\lambda}_v$							-0.473	0.060	0.212
							(0.362)	(0.341)	(0.353)
Number of Villages	25,740	25,739	25,739	25,740	25,739	25,739	25,740	25,739	25,739
Landholdings measure				**Wetland Holdings**					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pareto exponent $\widehat{\lambda}_v$	-0.018	0.050	0.031	-0.018	0.047	0.031	0.009	-0.092	-0.044
	(0.014)	(0.017)**	(0.017)**	(0.016)	(0.018)**	(0.018)**	(0.024)	(0.033)**	(0.028)*
Δ rainfall shock	0.042	0.234	0.133	0.049	0.165	0.055	0.039	0.296	0.113
	(0.139)	(0.150)	(0.140)	(0.174)	(0.186)	(0.177)	(0.172)	(0.183)*	(0.173)
Δ rice price shock	-0.213	0.505	0.083	-0.213	0.498	0.083	0.328	-1.992	-1.154
	(0.437)	(0.498)	(0.459)	(0.437)	(0.499)	(0.459)	(0.593)	(0.721)*	(0.684)
Δ agri. GDP/capita shock	0.197	1.508	1.015	0.197	1.486	1.006	0.199	1.365	0.960
	(0.129)	(0.173)***	(0.154)***	(0.129)	(0.173)***	(0.153)***	(0.129)	(0.168)***	(0.151)***
Δ rainfall shock $\times\ \widehat{\lambda}_v$				-0.004	0.039	0.047	-0.003	-0.038	0.014
				(0.053)	(0.058)	(0.059)	(0.051)	(0.053)	(0.051)
Δ rice price shock $\times\ \widehat{\lambda}_v$							-0.297	1.355	0.678
							(0.236)	(0.346)***	(0.307)**
Observations	23,855	23,852	23,852	23,855	23,852	23,852	23,855	23,852	23,852
Landholdings measure				**Paddy Area Planted, 2002-3**					
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Pareto exponent $\widehat{\lambda}_v$	0.015	0.030	0.034	0.015	0.016	0.033	0.029	-0.080	-0.059
	(0.017)	(0.017)	(0.017)*	(0.017)	(0.017)	(0.018)**	(0.032)	(0.049)**	(0.042)**
Δ rainfall shock	0.110	0.354	0.228	-0.075	0.103	0.034	-0.068	0.160	0.103
	(0.141)	(0.152)**	(0.143)*	(0.179)	(0.189)	(0.183)	(0.181)	(0.191)	(0.187)
Δ rice price shock	-0.250	0.196	-0.102	-0.238	0.192	-0.090	0.007	-1.398	-1.652
	(0.430)	(0.485)	(0.454)	(0.433)	(0.483)	(0.454)	(0.692)	(0.856)	(0.796)**
Δ agri. GDP/capita shock	0.110	1.164	0.740	0.110	1.122	0.721	0.110	1.111	0.666
	(0.131)	(0.167)***	(0.153)***	(0.130)	(0.167)***	(0.153)***	(0.130)	(0.165)***	(0.149)***
Δ rainfall shock $\times\ \widehat{\lambda}_v$				0.114	0.149	0.117	0.109	0.122	0.078
				(0.057)**	(0.064)**	(0.061)*	(0.058)*	(0.064)*	(0.064)
Δ rice price shock $\times\ \widehat{\lambda}_v$							-0.151	0.916	0.943
							(0.298)	(0.444)**	(0.392)***
Observations	24,158	24,157	24,157	24,158	24,157	24,157	24,158	24,157	24,157

Notes: The table reports estimates analogous to those in Tables 5 and 6 in the paper but augmented with a variable capturing overall agricultural GDP shocks as described in Section 1.9.4. The dependent variable in all specifications is $\Delta \ln M_{v,t+1}/N_{v,t+1}$ and has mean 0.11. Standard errors are clustered at the district level and significance levels are based on a block bootstrap$-t$ procedure. Additional covariates in all specifications but not reported here include all those reported under Table 5 in the paper. See Appendix B for details on the semiparametric (Semipar.) and parametric (Param.) correction procedures.

Table 1.16: Breaking Out Rainfall Shocks in Periods t and $t-1$

Correction Procedure	None	Semiparametric		Parametric
1st Stage Estimator	None	SU-LPM	SNP-ML	BiProbit
	(1)	(2)	(3)	(4)
Δ rainfall shock	0.098	0.415	0.572	0.296
	(0.127)	(0.133)***	(0.156)***	(0.128)**
	(5)	(6)	(7)	(8)
rainfall shock, t	0.159	0.407	0.552	0.309
	(0.127)	(0.132)***	(0.153)***	(0.126)**
rainfall shock, $t-1$	-0.111	-0.415	-0.628	-0.297
	(0.124)	(0.133)***	(0.154)***	(0.127)**
Number of Villages	26,529	26,527	26,527	26,527

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The top panel, reproduced from Table 5, takes the difference in cumulative log rainfall deviations between periods t (2006-8) and $t-1$ (2003-5). The bottom panel allows cumulative log rainfall deviations in periods t and $t-1$ to enter separately. The dependent variable in all specifications is $\Delta \ln M_{v,t+1}/N_{v,t+1}$ and has mean 0.11. Standard errors are clustered at the district level and significance levels are based on a block bootstrap$-t$ procedure. Additional covariates in all specifications but not reported here include all those reported under Table 5 in the paper.

Table 1.17: Full Elaboration of Annual Rainfall Shocks

Correction Procedure	None	Semipar.	Param.	None	Semipar.	Param.
1st Stage Estimator	None	SU-LPM	BiProbit	None	SU-LPM	BiProbit
	(1)	(2)	(3)	(4)	(5)	(6)
Pareto exponent $\widehat{\lambda}_v$	-0.004	0.053	0.040	-0.008	-0.056	-0.046
	(0.017)	(0.017)**	(0.018)**	(0.034)	(0.034)	(0.033)
θ_{a3}: log rainfall deviation, 2003	0.049	-0.662	-0.315	-0.011	0.230	0.063
	(0.213)	(0.228)	(0.220)	(0.322)	(0.331)	(0.322)
θ_{a4}: log rainfall deviation, 2004	0.424	2.475	1.379	0.649	2.319	1.568
	(0.441)	(0.466)**	(0.467)**	(0.680)	(0.660)**	(0.641)**
θ_{a5}: log rainfall deviation, 2005	-0.709	-3.063	-1.839	-0.701	-2.803	-1.729
	(0.201)***	(0.298)***	(0.238)***	(0.288)**	(0.362)***	(0.303)***
θ_{a6}: log rainfall deviation, 2006	0.098	0.674	0.135	-0.211	-0.830	-1.150
	(0.394)	(0.403)	(0.401)	(0.537)	(0.541)	(0.531)*
θ_{a7}: log rainfall deviation, 2007	0.636	-1.680	-0.662	0.613	-0.775	0.135
	(0.454)	(0.547)	(0.510)	(0.610)	(0.665)	(0.640)
θ_{a8}: log rainfall deviation, 2008	-0.452	2.101	1.267	-0.579	0.888	0.501
	(0.380)	(0.461)**	(0.449)**	(0.623)	(0.674)	(0.659)
$\theta_{a3\lambda}$: log rainfall deviation, 2003 $\times$ $\widehat{\lambda}_v$				0.043	-0.580	-0.243
				(0.169)	(0.174)**	(0.169)
$\theta_{a4\lambda}$: log rainfall deviation, 2004 $\times$ $\widehat{\lambda}_v$				-0.149	0.100	-0.108
				(0.342)	(0.337)	(0.323)
$\theta_{a5\lambda}$: log rainfall deviation, 2005 $\times$ $\widehat{\lambda}_v$				0.001	-0.167	-0.073
				(0.121)	(0.145)	(0.119)
$\theta_{a6\lambda}$: log rainfall deviation, 2006 $\times$ $\widehat{\lambda}_v$				0.217	1.037	0.865
				(0.271)	(0.277)***	(0.268)***
$\theta_{a7\lambda}$: log rainfall deviation, 2007 $\times$ $\widehat{\lambda}_v$				-0.014	-0.654	-0.596
				(0.265)	(0.279)*	(0.269)**
$\theta_{a8\lambda}$: log rainfall deviation, 2008 $\times$ $\widehat{\lambda}_v$				0.091	0.794	0.534
				(0.348)	(0.370)**	(0.360)*
$\boldsymbol{\sum_{s=3}^{5} \theta_{as}}$	**-0.236**	**-1.250**	**-0.775**	**-0.063**	**-0.253**	**-0.099**
$\boldsymbol{H_0}$: $\boldsymbol{\sum_{s=3}^{5} \theta_{as} = 0}$ **[p-value]**	**[0.571]**	**[0.003]**	**[0.080]**	**[0.914]**	**[0.655]**	**[0.859]**
$\boldsymbol{\sum_{s=6}^{8} \theta_{as}}$	**0.283**	**1.094**	**0.740**	**-0.177**	**-0.717**	**-0.514**
$\boldsymbol{H_0}$: $\boldsymbol{\sum_{s=6}^{8} \theta_{as} = 0}$ **[p-value]**	**[0.521]**	**[0.011]**	**[0.103]**	**[0.793]**	**[0.271]**	**[0.425]**
$\boldsymbol{\sum_{s=3}^{8} \theta_{as}}$	**0.046**	**-0.155**	**-0.035**	**-0.240**	**-0.971**	**-0.612**
$\boldsymbol{H_0}$: $\boldsymbol{\sum_{s=3}^{8} \theta_{as} = 0}$ **[p-value]**	**[0.632]**	**[0.141]**	**[0.733]**	**[0.268]**	**[0.0001]**	**[0.003]**
$\boldsymbol{\sum_{s=3}^{5} \theta_{as\lambda}}$				**-0.105**	**-0.647**	**-0.424**
$\boldsymbol{H_0}$: $\boldsymbol{\sum_{s=3}^{5} \theta_{as\lambda} = 0}$ **[p-value]**				**[0.699]**	**[0.024]**	**[0.103]**
$\boldsymbol{\sum_{s=6}^{8} \theta_{as\lambda}}$				**0.293**	**1.178**	**0.804**
$\boldsymbol{H_0}$: $\boldsymbol{\sum_{s=6}^{8} \theta_{as\lambda} = 0}$ **[p-value]**				**[0.369]**	**[0.001]**	**[0.011]**
$\boldsymbol{\sum_{s=3}^{8} \theta_{as\lambda}}$				**0.189**	**0.531**	**0.379**
$\boldsymbol{H_0}$: $\boldsymbol{\sum_{s=3}^{8} \theta_{as\lambda} = 0}$ **[p-value]**				**[0.119]**	**[$<$ 0.001]**	**[0.003]**
Number of Villages	26,529	26,527	26,527	26,529	26,527	26,527

Notes: The table reports estimates of equation (13) (in columns 1-3) and (14) (in columns 4-6) in the text with a fully elaborated set of annual rainfall shocks instead of cumulating those shocks over three seasons into a single rainfall shock term. Standard errors are clustered at the district level and significance levels are based on a block bootstrap$-t$ procedure. The dependent variable in all specifications is $\Delta \ln M_{v,t+1}/N_{v,t+1}$ and has mean 0.11. Additional covariates in all specifications but not reported here include all those reported under Table 5 in the paper. The p-values in the bottom panel are based on F tests.

Table 1.18: Alternative Rice Price Shock Measures

Correction Procedure	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.
1st Stage Estimator	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit
Price Shock Proxy	Δ*annualized log rice price growth*				Δ*log average rice price*				*log shipping distance pass-through model*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Pareto exponent $\widehat{\lambda}_v$	0.072	0.050	-0.113	-0.070	0.058	0.043	-0.338	-0.118	0.070	0.049	0.933	1.640
	(0.018)***	(0.017)***	(0.030)***	(0.030)**	(0.019)**	(0.017)**	(0.128)**	(0.111)	(0.019)***	(0.018)***	(0.842)	(0.801)*
Δ price shock	1.287	0.625	-2.234	-1.503								
	(0.502)*	(0.451)	(0.709)*	(0.688)**								
Δ price shock $\times$ $\widehat{\lambda}_v$			1.913	1.155								
			(0.335)***	(0.327)***								
Δ avg. price					1.630	0.396	-0.311	-0.269				
					(0.613)**	(0.524)	(0.688)	(0.647)				
Δ avg. price $\times$ $\widehat{\lambda}_v$							1.158	0.442				
							(0.377)***	(0.320)**				
log shipping distance to THA/VNM									-0.852	-0.866	-0.718	-0.529
									(0.319)	(0.295)**	(0.383)	(0.357)
$\times$ $\widehat{\lambda}_v$											-0.041	-0.172
											(0.114)	(0.106)
log distance to nearest port									-0.116	-0.001	0.041	0.064
									(0.053)	(0.050)	(0.065)	(0.062)
$\times$ $\widehat{\lambda}_v$											-0.098	-0.042
											(0.023)**	(0.022)
Number of villages	24,537	24,537	24,537	24,537	24,537	24,537	24,537	24,537	24,537	24,537	24,537	24,537

Notes: Columns 1-4 are reproduced from Tables 5 and 6 in the paper. Columns 5-8 uses the log difference in average rice prices between periods as the measure of the "shock". Columns 9-12 apply the insights from the trade model in Appendix 1.9.5 to use a distance-based proxy for the local intensity of the price shock. The dependent variable in all specifications is $\Delta \ln M_{v,t+1}/N_{v,t+1}$ and has mean 0.11. Standard errors are clustered at the district level and significance levels are based on a block bootstrap$-t$ procedure. Additional covariates in all specifications but not reported here include all those reported under Table 5 in the paper. See Appendix B in the paper for details on the semiparametric (Semipar.) and parametric (Param.) correction procedures.

Table 1.19: Alternative Choices of $\underline{R}$ in Estimating λ_v

Correction Procedure	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.
1st Stage Estimator	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	$\underline{R} = 0.1$ Hectares						**$\underline{R} = 0.15$ Hectares**					
Pareto exponent $\widehat{\lambda}_v$	0.072	0.050	0.073	0.052	-0.113	-0.070	0.057	0.041	0.053	0.040	-0.090	-0.057
	$(0.018)^{***}$	$(0.017)^{***}$	$(0.018)^{***}$	$(0.017)^{***}$	$(0.030)^{***}$	$(0.030)^{**}$	$(0.015)^{***}$	$(0.014)^{***}$	$(0.015)^{***}$	$(0.014)^{***}$	$(0.023)^{***}$	$(0.026)^{**}$
Δ rainfall shock	0.305	0.212	0.162	0.069	0.262	0.135	0.320	0.227	0.270	0.137	0.357	0.200
	$(0.134)^{*}$	(0.130)	(0.167)	(0.166)	(0.168)	(0.161)	$(0.135)^{**}$	$(0.130)^{*}$	(0.167)	(0.161)	$(0.168)^{*}$	(0.159)
Δ rainfall shock $\times$ $\widehat{\lambda}_v$			0.082	0.085	0.028	0.048			0.025	0.047	-0.014	0.017
			(0.049)	$(0.058)^{*}$	(0.052)	(0.051)			(0.042)	(0.046)	(0.045)	(0.043)
Δ price shock	1.287	0.625	1.304	0.627	-2.234	-1.503	1.319	0.615	1.326	0.609	-1.654	-1.265
	$(0.502)^{*}$	(0.451)	$(0.503)^{*}$	(0.452)	$(0.709)^{*}$	$(0.688)^{*}$	$(0.503)^{*}$	(0.453)	$(0.503)^{*}$	(0.452)	(0.657)	(0.682)
Δ price shock $\times$ $\widehat{\lambda}_v$					1.913	1.155					1.451	0.908
					$(0.335)^{***}$	$(0.327)^{***}$					$(0.266)^{***}$	$(0.285)^{***}$
Number of Villages	26,527	26,527	26,527	26,529	26,527	26,527	26,435	26,435	26,435	26,435	26,435	26,435
	$\underline{R} = 0.2$ Hectares						**$\underline{R} = 0.25$ Hectares**					
Pareto exponent $\widehat{\lambda}_v$	0.034	0.036	0.030	0.034	-0.093	-0.057	0.026	0.025	0.024	0.027	-0.072	-0.065
	$(0.013)^{*}$	$(0.014)^{***}$	$(0.012)^{**}$	$(0.014)^{***}$	$(0.023)^{***}$	$(0.030)^{**}$	$(0.012)^{*}$	$(0.012)^{**}$	$(0.012)^{**}$	$(0.011)^{**}$	$(0.022)^{***}$	$(0.029)^{***}$
Δ rainfall shock	0.291	0.204	0.219	0.056	0.299	0.145	0.326	0.225	0.161	-0.000	0.288	0.116
	$(0.136)^{*}$	(0.131)	(0.162)	(0.153)	(0.164)	(0.153)	$(0.136)^{*}$	$(0.131)^{*}$	(0.165)	(0.154)	$(0.167)^{*}$	(0.158)
Δ rainfall shock $\times$ $\widehat{\lambda}_v$			0.034	0.072	-0.000	0.033			0.074	0.106	0.023	0.055
			(0.036)	$(0.039)^{**}$	(0.037)	(0.037)			$(0.037)^{*}$	$(0.037)^{***}$	(0.039)	(0.038)
Δ price shock	1.338	0.627	1.331	0.622	-1.361	-1.188	1.369	0.638	1.358	0.637	-0.998	-1.331
	$(0.498)^{*}$	(0.447)	$(0.497)^{*}$	(0.447)	(0.661)	(0.731)	$(0.499)^{*}$	(0.444)	$(0.498)^{**}$	(0.444)	(0.651)	$(0.735)^{*}$
Δ price shock $\times$ $\widehat{\lambda}_v$					1.235	0.819					0.985	0.825
					$(0.248)^{***}$	$(0.289)^{***}$					$(0.229)^{***}$	$(0.274)^{***}$
Number of Villages	26,346	26,346	26,346	26,346	26,346	26,346	26,242	26,242	26,242	26,242	26,242	26,242

Notes: The table reports estimates analogous to those in Tables 5 and 6 in the paper but allowing for alternative $\underline{R}$ thresholds in the estimation of λ_v (and the share of households above $\underline{R}$). Baseline estimates using $\underline{R} = 0.1$ Ha are reported in columns 1-6 of the top panel. All estimates in the table are for wetland holdings. The dependent variable in all specifications is $\Delta \ln M_{v,t+1}/N_{v,t+1}$ and has mean 0.11. Standard errors are clustered at the district level, and significance levels are based on a block bootstrap$-t$ procedure. Additional covariates in all specifications but not reported here include all those reported under Table 5 in the paper. See Appendix B in the paper for details on the semiparametric (Semipar.) and parametric (Param.) correction procedures.

Table 1.20: Parsimony and Parallel Trends in Estimates of θ_a and θ_p

Correction Procedure	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.
1st Stage Estimator	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit
Variables Removed from Baseline	—		Skill Distribution		. . . Migration Costs	
	(1)	(2)	(3)	(4)	(5)	(6)
Pareto exponent $\widehat{\lambda}_v$	0.072	0.050	0.073	0.050	0.073	0.049
	$(0.018)^{***}$	$(0.017)^{***}$	$(0.018)^{***}$	$(0.017)^{***}$	$(0.018)^{***}$	$(0.017)^{**}$
Δ rainfall shock	0.305	0.212	0.304	0.212	0.327	0.209
	$(0.134)^{**}$	(0.130)	$(0.137)^{*}$	(0.133)	$(0.138)^{**}$	$(0.133)^{*}$
Δ price shock	1.287	0.625	1.257	0.651	1.340	0.529
	$(0.502)^{*}$	(0.451)	(0.498)	(0.450)	$(0.457)^{*}$	(0.417)
Variables Removed from Baseline	. . . Land Distribution		. . . Price Shock		. . . Rainfall Shock	
	(7)	(8)	(9)	(10)	(11)	(12)
Δ rainfall shock	0.332	0.206	0.319	0.207		
	$(0.141)^{**}$	(0.133)	$(0.141)^{*}$	(0.134)		
Δ price shock	1.146	0.361			1.593	0.574
	$(0.452)^{*}$	(0.415)			$(0.487)^{**}$	(0.424)
Number of Villages	24,537	24,537	24,537	24,537	24,537	24,537

Notes: The table reports estimates of equation (13) based on progressively more parsimonious specifications than the baseline reported in columns 1-2, reproduced from Table 5. All estimates in the table are based on $\widehat{\lambda}_v$ using wetland holdings (results are similar for other measures of landholdings). The dependent variable in all specifications is $\Delta \ln M_{v,t+1}/N_{v,t+1}$ and has mean 0.11. Standard errors are clustered at the district level, and significance levels are based on a block bootstrap$-t$ procedure. Additional covariates in all specifications but not reported here include all those reported in Table 5 in the paper. See Appendix B in the paper for details on the semiparametric (Semipar.) and parametric (Param.) correction procedures.

Table 1.21: Parsimony and Parallel Trends in Estimates of $\theta_{a\lambda}$ and $\theta_{p\lambda}$

Correction Procedure	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.
1st Stage Estimator	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit
Variables Removed from Baseline	—		Skill Distribution		. . . Migration Costs	
	(1)	(2)	(3)	(4)	(5)	(6)
Pareto exponent $\widehat{\lambda}_v$	0.073	0.052	0.073	0.052	0.073	0.052
	(0.018)***	(0.017)***	(0.019)***	(0.018)***	(0.019)***	(0.018)***
Δ rainfall shock	0.082	0.085	0.084	0.086	0.084	0.086
	(0.049)	(0.058)*	(0.049)	(0.058)	(0.049)	(0.058)
Δ rainfall shock $\times$ $\widehat{\lambda}_v$	0.162	0.069	0.157	0.067	0.157	0.067
	(0.167)	(0.166)	(0.171)	(0.169)	(0.171)	(0.169)
Δ price shock	1.304	0.627	1.276	0.655	1.276	0.655
	(0.503)*	(0.452)	(0.499)*	(0.450)	(0.499)**	(0.450)
	(7)	(8)	(9)	(10)	(11)	(12)
Pareto exponent $\widehat{\lambda}_v$	-0.113	-0.070	-0.112	-0.070	-0.122	-0.078
	(0.030)***	(0.030)**	(0.030)**	(0.030)***	(0.030)***	(0.030)***
Δ rainfall shock	0.028	0.048	0.033	0.050	0.043	0.067
	(0.052)	(0.051)	(0.052)	(0.051)	(0.052)	(0.051)
Δ rainfall shock $\times$ $\widehat{\lambda}_v$	0.262	0.135	0.253	0.131	0.259	0.100
	(0.168)	(0.161)	(0.171)	(0.164)	(0.171)	(0.164)
Δ price shock	-2.234	-1.503	-2.242	-1.493	-2.345	-1.759
	(0.709)*	(0.688)*	(0.712)*	(0.692)*	(0.674)*	(0.680)**
Δ price shock $\times$ $\widehat{\lambda}_v$	1.913	1.155	1.908	1.164	2.016	1.247
	(0.335)***	(0.327)***	(0.340)***	(0.331)***	(0.341)***	(0.336)***
Number of Villages	24,537	24,537	24,537	24,537	24,537	24,537

Notes: The table considers progressively more parsimonious specifications than the baseline reported in columns 1-2, reproduced from Table 6. See the Notes to Table 1.20 for further details.

Table 1.22: Controlling for Omitted Demographic Variables

Correction Procedure	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.	Semipar.	Param.
1st Stage Estimator	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit	SU-LPM	BiProbit
Landholdings measure:	**Total Agricultural Landholdings**				**Wetland Holdings**				**Paddy Planted, 2002**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Pareto exponent $\widehat{\lambda}_v$	0.039	0.037	0.005	0.004	0.073	0.048	-0.084	-0.044	0.044	0.048	-0.057	-0.053
	(0.017)*	(0.018)*	(0.034)	(0.034)	(0.017)***	(0.017)***	(0.028)**	(0.027)	(0.016)**	(0.016)**	(0.043)	(0.039)*
Δ rainfall shock	0.392	0.299	0.199	0.188	0.266	0.201	0.195	0.120	0.352	0.248	-0.522	-1.238
	(0.134)**	(0.127)**	(0.169)	(0.162)	(0.134)*	(0.129)	(0.167)	(0.160)	(0.141)**	(0.135)*	(0.797)	(0.745)*
Δ price shock	0.493	0.333	0.318	-0.168	1.321	0.609	-1.628	-1.005	0.937	0.450	0.137	0.106
	(0.462)	(0.437)	(0.690)	(0.679)	(0.510)*	(0.464)	(0.694)	(0.651)	(0.497)	(0.455)	(0.174)	(0.173)
Δ rainfall shock $\times\ \widehat{\lambda}_v$			0.120	0.074			0.042	0.050			0.132	0.086
			(0.072)**	(0.071)			(0.051)	(0.049)			(0.060)**	(0.060)
Δ price shock $\times\ \widehat{\lambda}_v$			0.106	0.327			1.585	0.872			0.823	1.001
			(0.326)	(0.331)			(0.312)***	(0.294)***			(0.392)**	(0.370)***
average household size, 2000	-0.003	-0.014	-0.003	-0.014	-0.006	-0.023	-0.007	-0.023	-0.009	-0.022	-0.010	-0.020
	(0.023)	(0.022)	(0.023)	(0.022)	(0.025)	(0.024)	(0.024)	(0.024)	(0.024)	(0.023)	(0.024)	(0.022)
15-29 year old population share, 2000	0.220	0.239	0.263	0.261	0.528	0.698	0.618	0.743	0.385	0.520	0.434	0.583
	(0.301)	(0.303)	(0.301)	(0.303)	(0.300)	(0.298)	(0.297)	(0.297)*	(0.304)	(0.299)	(0.301)	(0.297)
internal migrant share, 2000	0.655	0.594	0.662	0.597	1.068	0.821	0.978	0.785	0.864	0.705	0.841	0.665
	(0.159)***	(0.164)***	(0.159)***	(0.163)***	(0.157)***	(0.165)***	(0.156)***	(0.162)***	(0.164)***	(0.171)***	(0.164)***	(0.171)***
Number of Villages	26,527	26,527	26,527	26,527	24,537	24,537	24,537	24,537	24,476	24,476	24,476	24,476

Notes: The table augments the baseline estimates in Tables 5 and 6 of the paper with the mean of average household size in the village, the share of the population aged 15-29 in 2000, and the share of individuals that resided in a different district in 1995—all of which are obtained from the Population Census of 2000. The dependent variable in all specifications is $\Delta \ln M_{v,t+1}/N_{v,t+1}$ and has mean 0.11. Standard errors are clustered at the district level, and significance levels are based on a block bootstrap$-t$ procedure. Additional covariates in all specifications but not reported here include all those reported under Table 5 in the paper. See Appendix B in the paper for details on the semiparametric (Semipar.) and parametric (Param.) correction procedures.

Table 1.23: Accounting for Measurement and Reporting Outliers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Landholdings measure	**Total Agricultural Landholdings**						
Pareto exponent $\widehat{\lambda}_v$	0.039	0.036	0.058	0.037	0.076	0.091	-0.029
	(0.017)*	(0.017)	(0.024)	(0.017)	(0.019)**	(0.025)**	(0.021)
Δ price shock	0.409	0.430	0.394	0.378	0.425	0.386	0.781
	(0.448)	(0.447)	(0.449)	(0.446)	(0.491)	(0.490)	(0.482)
Δ rainfall shock	0.415	0.403	0.415	0.420	0.368	0.369	0.575
	(0.133)**	(0.133)**	(0.134)***	(0.133)***	(0.138)*	(0.139)**	(0.170)***
Number of Villages	26,527	26,527	26,294	26,482	23,539	23,296	19,031
Landholdings measure	**Wetland Holdings**						
Pareto exponent $\widehat{\lambda}_v$	0.072	0.069	0.086	0.071	0.109	0.104	0.017
	(0.018)***	(0.018)***	(0.022)***	(0.018)***	(0.021)***	(0.025)***	(0.020)
Δ price shock	1.287	1.321	1.256	1.256	1.299	1.309	1.378
	(0.502)*	(0.500)*	(0.496)	(0.499)*	(0.532)*	(0.535)**	(0.548)*
Δ rainfall shock	0.305	0.293	0.328	0.311	0.307	0.315	0.445
	(0.134)**	(0.134)*	(0.135)*	(0.134)**	(0.141)*	(0.143)*	(0.172)**
Number of Villages	24,537	24,537	24,304	24,493	21,929	21,705	17,286
	Paddy Planted, 2002						
Pareto exponent $\widehat{\lambda}_v$	0.043	0.041	0.065	0.044	0.074	0.088	0.038
	(0.016)**	(0.016)*	(0.021)**	(0.016)**	(0.019)***	(0.024)**	(0.018)
Δ price shock	0.919	0.955	0.989	0.914	0.820	0.945	1.294
	(0.487)	(0.485)	(0.489)	(0.485)	(0.523)	(0.525)	(0.527)*
Δ rainfall shock	0.390	0.376	0.398	0.393	0.377	0.379	0.541
	(0.139)**	(0.140)**	(0.139)**	(0.140)**	(0.148)**	(0.149)**	(0.179)***
Number of Villages	24,855	24,855	24,650	24,812	22,136	21,924	17,615
Reporting Frequency Indicators		Yes				Yes	
Lambda Trimmed			Yes			Yes	
Migration Reporting Outliers Removed				Yes		Yes	
Perfect Match Stage					Yes	Yes	
High Illegal Migration Provinces Removed							Yes

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; Standard errors are clustered at the district level in all specifications, and significance levels are based on a block bootstrap$-t$ procedure. Column 1 reproduces the baseline estimates from Table 5; column 2 includes five indicators for the frequency of population register updating in the village (see Appendix C in the paper); column 3 trims the bottom 1 and top 99 percentiles of the distribution of $\widehat{\lambda}_v$; column 4 removes those villages for which the reporting format of *Podes* 2005 and/or 2008 results in top-censoring of migrant stocks in certain villages; column 5 retains only those villages for which *Podes* 2005 and/or 2008 could be matched exactly on administrative codes and village name (see Appendix 1.9.8); column 6 combines the previous four restrictions; column 7 drops villages in East Java, West Nusa Tenggara and provinces in Kalimantan, all of which are conjectured to have high illegal emigration outflows according to Bank Indonesia (2009). Additional covariates in all specifications but not reported here include all those reported under Table 5 in the paper or mentioned in the notes therein. All estimates are based on the semiparametric correction procedure (see Appendix B in the paper).

Table 1.24: Accounting for Outliers in Estimating $\theta_{a\lambda}$ and $\theta_{p\lambda}$

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Land-holdings measure	**Total Agricultural Land-holdings**						
Pareto exponent $\widehat{\lambda}_v$	-0.007	-0.011	-0.038	-0.012	0.063	0.016	-0.032
	(0.035)	(0.035)	(0.055)	(0.035)	(0.044)	(0.060)	(0.041)
Δ price shock	-0.016	-0.016	-0.779	-0.101	0.537	-0.270	0.741
	(0.688)	(0.687)	(0.909)	(0.689)	(0.808)	(0.932)	(0.769)
Δ rainfall shock	0.225	0.204	0.268	0.227	0.261	0.219	0.425
	(0.169)	(0.169)	(0.203)	(0.169)	(0.182)	(0.214)	(0.210)**
Δ price shock $\times$ $\widehat{\lambda}_v$	0.267	0.284	0.719	0.302	-0.089	0.347	0.041
	(0.329)	(0.329)	(0.520)	(0.332)	(0.409)	(0.540)	(0.380)
Δ rainfall shock $\times$ $\widehat{\lambda}_v$	0.119	0.126	0.062	0.122	0.061	0.055	0.102
	(0.074)**	(0.073)*	(0.107)	(0.074)**	(0.077)	(0.109)	(0.087)
Number of Villages	26,527	26,527	26,294	26,482	23,539	23,296	19,031
Land-holdings measure	**Wetland Holdings**						
Pareto exponent $\widehat{\lambda}_v$	-0.113	-0.113	-0.098	-0.121	-0.081	-0.072	-0.028
	(0.030)***	(0.030)**	(0.041)*	(0.029)***	(0.032)**	(0.045)	(0.035)
Δ price shock	-2.234	-2.132	-1.998	-2.360	-2.079	-1.672	0.145
	(0.709)*	(0.701)*	(0.833)	(0.705)*	(0.752)*	(0.899)	(0.780)
Δ rainfall shock	0.262	0.237	0.129	0.281	0.187	0.192	0.477
	(0.168)	(0.168)	(0.207)	(0.168)	(0.181)	(0.221)	(0.209)**
Δ price shock $\times$ $\widehat{\lambda}_v$	1.913	1.876	1.738	1.965	1.801	1.522	0.603
	(0.335)***	(0.333)***	(0.431)**	(0.335)***	(0.354)***	(0.463)***	(0.366)
Δ rainfall shock $\times$ $\widehat{\lambda}_v$	0.028	0.037	0.111	0.021	0.066	0.062	-0.043
	(0.052)	(0.052)	(0.084)	(0.051)	(0.056)	(0.091)	(0.055)
Number of Villages	24,537	24,537	24,304	24,493	21,929	21,705	17,286
	Paddy Planted, 2002						
Pareto exponent $\widehat{\lambda}_v$	-0.083	-0.087	-0.101	-0.082	-0.020	-0.056	-0.030
	(0.046)**	(0.044)**	(0.055)**	(0.046)*	(0.051)	(0.064)	(0.051)
Δ price shock	-1.031	-1.042	-1.937	-1.035	-0.465	-1.447	-0.306
	(0.822)	(0.804)	(0.899)	(0.824)	(0.915)	(1.050)	(0.889)
Δ rainfall shock	0.167	0.131	0.252	0.173	0.177	0.201	0.561
	(0.176)	(0.177)	(0.198)	(0.176)	(0.202)	(0.223)	(0.220)**
Δ price shock $\times$ $\widehat{\lambda}_v$	1.116	1.144	1.696	1.115	0.711	1.336	0.841
	(0.423)***	(0.411)***	(0.506)***	(0.426)**	(0.482)*	(0.605)**	(0.456)**
Δ rainfall shock $\times$ $\widehat{\lambda}_v$	0.140	0.155	0.086	0.137	0.121	0.098	-0.055
	(0.065)**	(0.065)**	(0.087)	(0.064)**	(0.081)*	(0.102)	(0.075)
Number of Villages	24,855	24,855	24,650	24,812	22,136	21,924	17,615
Reporting Frequency Indicators		Yes				Yes	
Lambda Trimmed			Yes			Yes	
Migration Reporting Outliers Removed				Yes		Yes	
Perfect Match Stage					Yes	Yes	
High Illegal Migration Provinces Removed							Yes

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; Standard errors are clustered at the district level in all specifications, and significance levels are based on a block bootstrap$-t$ procedure. All estimates are based on the semiparametric correction procedure (see Appendix B in the paper). See the Notes to Table 1.23.

1.9.5 Further Details on Agricultural Income Shocks

In this section, I provide further background on the rice price shock subsequent to the ban on imports in 2004 as well as details on the time series properties of rainfall and rice prices.

Spatial Variation in the Rice Price Shock: A Simple Model

To understand how the import ban exerted differential pressure on local prices across regions, I first consider a simple model which micro-founds local rice prices based on the domestic market structure, imports, and world prices. The primary contribution of the model is to rationalize the lack of spatial arbitrage evident in Figure 7 in the paper.[82] I adapt the formulation for changes in national rice prices given in Warr (2008) to a model in which key parameters are allowed to vary across regions of the country. An important background assumption in my adapted model is that there are no strategic interactions among producers or consumers across villages, but local market power (among farmers) is possible in the sense of monopolistic competition. In Warr's model, there are no village subscripts v. For simplicity, I ignore the cross-village price elasticities such that changes in supply or demand conditions in village j have no effect on prices in village k.[83]

A key prediction of the model is that changes in rice prices vary across villages according to a simple expression relating proportional log changes in farmgate rice prices in village v in year t, p_v^d (d for domestic), to log changes in world

[82]The delayed effect of the import ban evident in that figure has a straightforward explanation. Imported rice was especially important in the months around harvests at the end of growing seasons with particularly low rainfall. Because the spring 2004 harvest occurred after a season of high rainfall, the lack of imported rice in early 2004 had little effect on prices. In fact, it was not until just prior to the primary harvest in spring 2005 after a season of low rainfall in certain regions that the lack of imports proved important as domestic rice prices began to escalate across Indonesia.

[83]Wimanda (2009) finds that the average speed of convergence for perishable goods across large regions of the country is approximately 9 months, which suggests that price deviations across markets and villages *within the same region (e.g., province)* should be even more short-lived. In short, though cross-village price elasticities exist, they are likely to be vanishingly small over the relatively long time horizons considered here.

prices, p^m, (m for imported)

$$\widehat{p}^d_{vt} = \epsilon^m_v \widehat{p}^m_t + \epsilon^o_v \widehat{p}^o_t \tag{1.26}$$

where ϵ^m_v is the passthrough elasticity from world prices, and ϵ^o_v is the elasticity of domestic rice prices with respect to changes in prices p^o of an index of other goods consumed within Indonesia.[84] The partial equilibrium form of the village v passthrough elasticity ϵ^m_v is given by

$$\epsilon^m_v = S^m_v(\rho_v + \eta_v)/(\chi^d_v + \rho_v S^m_v - \eta_v S^d_v), \tag{1.27}$$

where $\eta_v \leq 0$ is the overall price elasticity of demand for rice (composite of domestic and imported) in the geographically delineated markets relevant to village v; S^m_v is the share of imported rice in total rice expenditures and $S^d_v = 1 - S^m_v$ is the expenditure share on domestically-produced rice; ρ_v is the Armington elasticity of substitution between domestic and imported rice; and χ^d_v is the elasticity of domestic supply with respect to prices p^d.

As world prices declined from 2005 to late 2007 (see Figure 1 in the paper), the model above suggests that, net of the effect of the change in other prices, domestic prices should also have fallen. Instead, the import ban effectively imposed $\epsilon^m_v = 0$ *for all* villages (see Figure 1.10 below).[85] Conditional on other determinants of rice prices, the relevant counterfactual setting would be one in which villages with $\epsilon^m_v > 0$ before the ban experience a decline in real rice prices while villages with $\epsilon^m_v \approx 0$—those with zero import penetration in local markets—experience no change at all. In other words, given the price stabilizing role of imports in villages

[84] One concern with this approach is that Indonesia's import level directly affects world prices. Although there is some time series evidence that world prices are increasing in Indonesian imports, it is unclear whether the relationship is causal or due to the effect of climate shocks throughout Southeast Asia which reduce output in major rice-exporting countries and also increase demand for imports in Indonesia. By all accounts, Indonesia remains a price-taker in the world rice market. Dawe (2008), for example, identifies an optimal ad valorem tariff of around 4 percent, which is essentially indistinguishable from free trade.

[85] Small import shipments in late 2007 were undertaken as part of a limited government-licensed procurement from Thailand and Vietnam to be distributed largely through the Raskin program which provides heavily subsidized rice to households below and just above the poverty line.

with $\epsilon_v^m > 0$ before the ban, the model implies that the import ban should cause larger price increases in villages with a higher passthrough elasticity,

$$\widehat{p}_{vt}^d\Big|_{\epsilon_v^m>0} > \widehat{p}_{v't}^d\Big|_{\epsilon_{v'}^m\approx 0} \quad , \tag{1.28}$$

while the counterfactual implies the opposite

$$\widehat{p}_{vt}^d\Big|_{\epsilon_v^m>0} \leq \widehat{p}_{v't}^d\Big|_{\epsilon_{v'}^m\approx 0} \quad . \tag{1.29}$$

The relevant empirical question, then, is what determines variation in ϵ_v^m across villages.

According to equation (1.27), the local intensity of world price passthrough is governed by four parameters: the share of imports in local rice consumption, the price elasticities of supply and demand, and the Armington elasticity of substitution between domestic and imported rice. The most relevant predictions are that ϵ_v^m should be decreasing in the local price elasticity of supply and increasing in the share of imported rice in the markets which purchase village v output.[86] The limited available estimates suggest that supply elasticities vary considerably across regions and land types—0.15 on Java, 0.4 in Sumatra, 1.25 in Sulawesi for wetland paddy, and dryland supply elasticities are approximately twice as large (Warr, 2005). Moreover, given prevailing transportation and trade costs, the local pre-ban import penetration ratio should be decreasing in (i) the distance to the nearest international port and major wholesale markets, and (ii) the shipping distance from

[86]There are two other predictions less relevant to the first order discussion here. First, ϵ_v^m is decreasing in the Armington elasticity of substitution between domestic and imported rice. This elasticity should be quite homogenous across the country and relatively high (Warr (2008) estimates around 5) given that nearly all Indonesian rice production is of the Indica type which is the predominant variety produced in Southeast Asia and traded on world markets (Dawe, 2008). Second, ϵ_v^m is increasing in the consumer price elasticity of demand for (all) rice in the regions relevant to village v. Estimates from the mid-1990s suggest that the price elasticity of demand is approximately -0.45 on average across all regions of Indonesia (Friedman and Levinsohn, 2001), though I have preliminary evidence that this estimate may have fallen considerably in recent years. Most of the variation in this estimate occurs within rather than across regions as the wealthy can more readily substitute away from rice staples when prices rise. In some of the Outer Islands, however, availability of cassava and other tubers allow greater substitution away from rice and hence higher demand elasticities.

the nearest port to Bangkok and Ho Chi Minh City, the two primary markets from which the majority of Indonesia's rice imports originate. Indonesia's unique geography generates substantial variation in these distance-driven components of ϵ_v^m.

Using the monthly consumer rice price index described in the paper, Table 1.25 demonstrates that the empirical changes in rice prices from 2002-2008 are consistent with the model sketched above. I control for lagged rainfall levels to account for local supply shocks, but the main proxy for ϵ_v^m is the log average shipping distance to Thailand and Vietnam via the nearest port city in Indonesia.[87] Regardless of the growth horizon on the left hand side (monthly, semi-annual, or annual), the primary takeaway is that after the import ban in January 2004, prices grew slower in Indonesian cities farther removed from the main rice exporter shipping routes in Southeast Asia. Before the ban, the opposite was true. Figure 1.11 graphically depicts this main finding, which is consistent with equations (1.28) and (1.29). As elaborated in the paper, the distinct lack of spatial arbitrage evident in these results can be explained on more fundamentally by the disruption of path-dependent, international buyer-seller networks after the import ban.

There are other possible explanations for spatial variation in the price shocks. For example, in the absence of imports, one might expect prices in the outer islands to be increasing in the distance to large wholesale markets in Java from where the majority of intra-island rice trade, including redistribution of international imports, originates. At the same time, enforcement of the import ban is likely to be stronger in larger ports closer to Jakarta. In smaller ports on the outer islands, government officials have more limited capacity to enforce federal policies banning the private import of rice. Conditional on shipping distance to Bangkok and Ho Chi Minh City, then, the change in prices might be increasing in (i) the distance to Jakarta and (ii) the size of the port from which the markets

[87] At present, all international port cities in Indonesia (on e-ships) are treated as equally likely to have imported rice prior to the import ban conditional on distance. Future research will incorporate information on the actual spatial distribution of imported rice prior to the ban using data on the total value and weight of rice imports for every port in the country from 2000-2008.

relevant to village v source imported rice. Empirically, (i) seems to be the dominant explanation. Absent information on the intra-national rice trade (which is not available), however, it is difficult to make further use of this stylized fact.

On Measuring Rice Prices

A few issues concerning the price indices deserve mention. First, while the price index is only available in 44 cities across Indonesia, these data points are arguably representative of the average regional prices faced by rice producers in nearby rural villages. Relative to prices in local rural markets, these measures should be (i) less affected by supply shocks in small groups of villages, and, (ii) more likely to capture the general equilibrium impact of the import ban. Second, farmgate prices are not available at the regional level. Nevertheless, results would likely be unchanged if farmgate prices were used instead. Figure 1.12 demonstrates the high correlation between farmgate, wholesale, and consumer prices over the period under study. Third, in some regions of Indonesia, up to 15% the price index is actually comprised of cassava and other tubers. This does not pose a problem here since prices of cassava and other tubers were stagnant over the period under study and hence should have little affect on the overall index. Precise figures available upon request.

Another approach to measuring rice price shocks would be to use mean unit values estimated over a representative sample of households within every district in Indonesia as recorded in *Susenas* expenditure modules in mid-2002, -2005, and -2008. By taking the average unit values across all households within a given district, one can smooth out the non-classical measurement errors implicit in individual unit values (see Deaton, 1997a). However, the added cross-sectional variation from this approach comes at the expense of picking up additional variation in prices attributable to a major rice subsidy program, known as *Raskin* or Rice for the Poor, which provides roughly 50-60 percent of the population in rural areas with access to rice at a fraction of market prices. This can be seen quite clearly in Table 1.26. The spatial variation in mean unit values could therefore be determined largely by the interregional variation in welfare levels, making it difficult to interpret the change in unit values as a producer price shock. An important advantage of the rice price index used in the empirical analysis is that it is unaffected by variation in household welfare.

Time Series Properties of Rainfall and Rice Prices

An important feature of rice prices is their approximate unit root properties. This is demonstrated in Figure 1.13 which plots the p-values from augmented Dickey and Fuller (1979) tests of the null hypothesis that the *domestic* rice price index in region c has a unit root (the different color dots correspond to alternative lag structures). Acknowledging that rice prices across Indonesian cities are not independent, I also apply the heterogeneous panel unit root tests of Im et al. (2003) and Fisher's meta-analytic test, and in both cases, I fail to reject reject the null hypothesis that rice prices follow a unit root in all cities. Recognizing further the possibility that the structural breaks in prices around late 2005 evident in Figure 7 in the paper might be mistaken for unit roots, I apply city-specific Zivot and Andrews (2002) unit root tests which allow for an endogenous break in both trends and intercepts. Doing so, I fail to reject the null of a unit root in 41 out of 44 cities. This finding is moreover robust to the removal of seasonal trends as well as cross-sectional demeaning to account for spatial correlation. The procedure identifies structural breaks between 2004m11 and 2006m4 for 85 percent of cities.

Whereas rice prices tend to follow a unit root, rainfall levels are serially uncorrelated across seasons. Considering seasonal rainfall levels at the district level (adjusted for province-specific growing seasons) going back to 1953, I cannot reject the null hypothesis of covariance stationarity for any Indonesian district.

Effect of Rainfall and Rice Price Shocks on Income and Wages

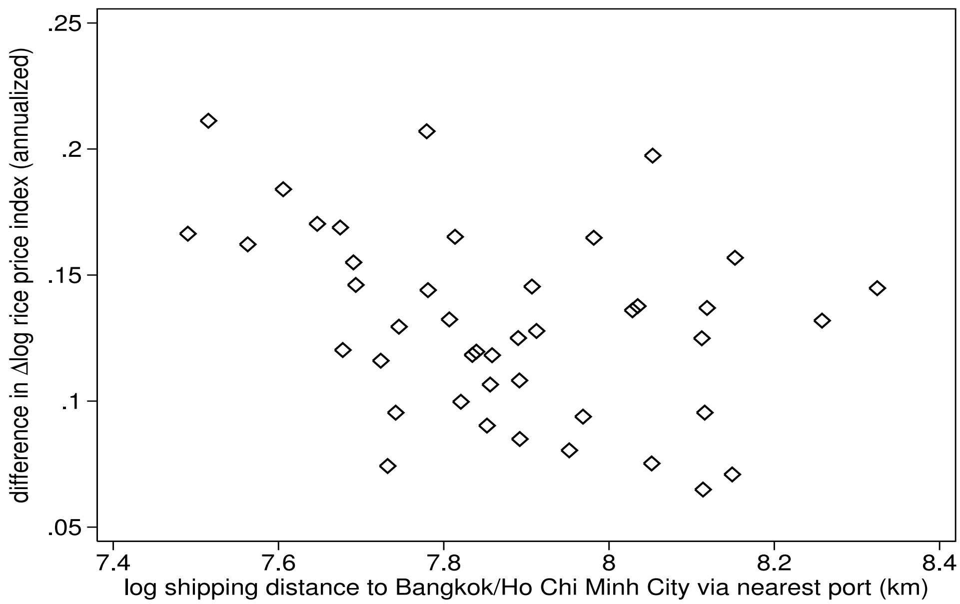
Rainfall has a strong positive relationship with time-varying agricultural productivity. Using a panel of district-level agricultural GDP from 2000-2010 (see

Section 1.9.4), I estimate an elasticity of agricultural GDP with respect to rainfall (in periods t and $t-1$) of around 0.15. This robust positive estimate is in line with results specific to rice output in Levine and Yang (2006) and Naylor et al. (2001). The relationship between rainfall and overall district-level GDP (i.e., across all sectors including both agricultural and non-agricultural), however, is null.

Figures

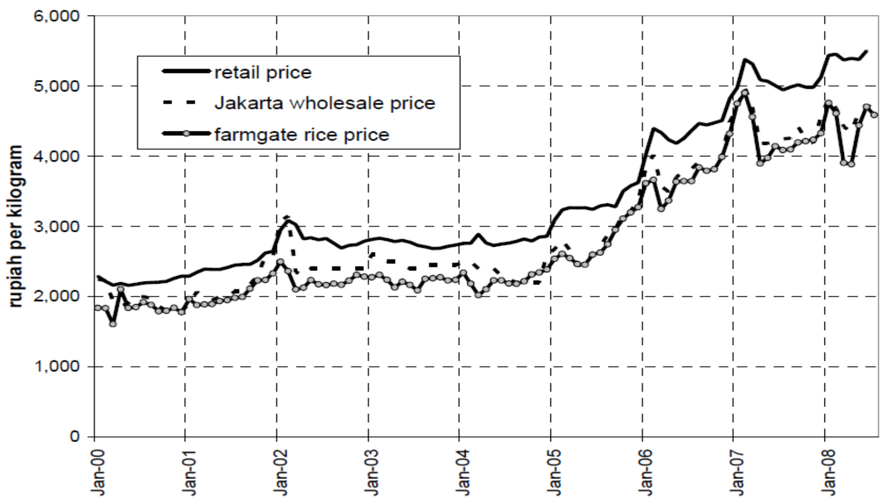
Notes: Data obtained from SITC Rev. 2 in the Comtrade-UN database on 5 December 2010. All categories of rice products are included in the figure. The uptick in 2007 is almost entirely due to emergency imports by the government logistical agency, Bulog.

Figure 1.10: Net Rice Imports in Indonesia, 1991-2008



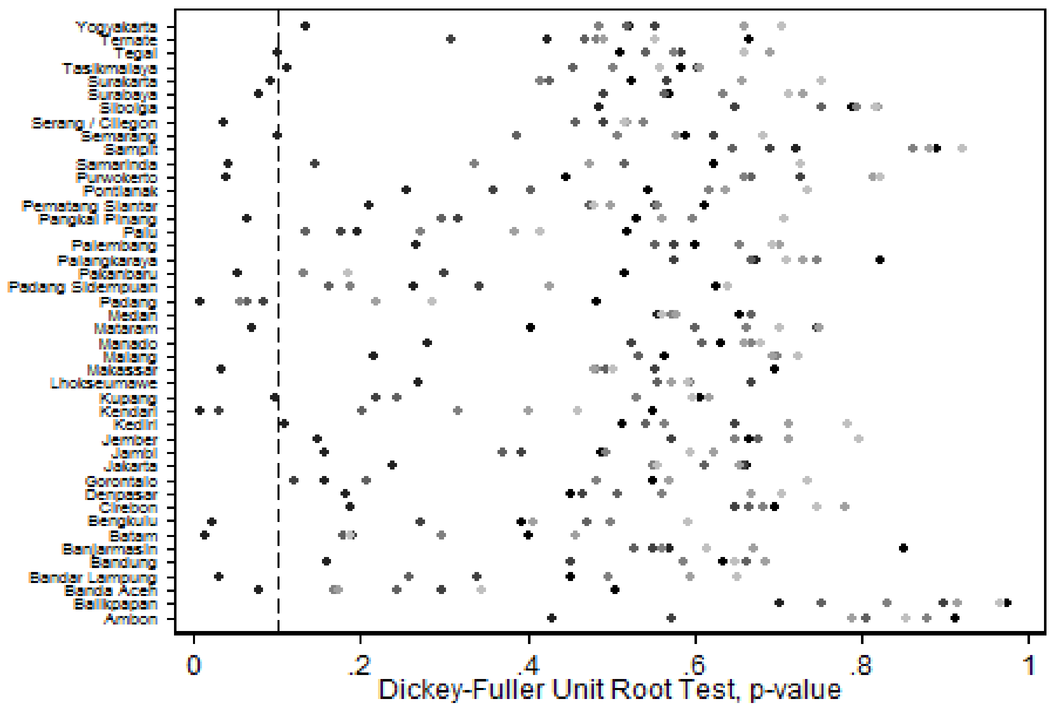
Notes: This figure demonstrates that rice prices grew faster in port cities closer to Bangkok and Ho Chi Minh City after the ban on rice imports. Monthly rice prices obtained from Wimanda (2009). Distances calculated as the sum of (i) the travel distance from the village to the district capital reported in *Podes* 2005, (ii) the great circle distance from the given Indonesian city to the nearest port, and (iii) the average shipping distance from the given Indonesian port to the port in Bangkok and Ho Chi Minh City. The port cities and shipping distances are obtained from http://e-ships.net. For cities with ports, I take the distance from the centroid of the city to the exact latitude/longitude of the port.

Figure 1.11: Spatial Variation in Rice Price Changes Around the Import Ban



Notes: Prices from January 2000 to January 2008. Farmgate price quoted in terms of wet paddy. After drying and milling, 100 kg of wet paddy produces 55 kg of rice. The figure is reproduced from Timmer (2008), but the original source is Peter B. Rosner from Ministry of Trade and Central Bureau of Statistics (BPS).

Figure 1.12: Nominal Wholesale/Retail/Farmgate Prices (monthly 2000-2008)



Notes: The monthly rice price index is from Wimanda (2009). The circles indicate p-values from augmented Dickey-Fuller unit roots for each of the 44 cities where the colors are lightest for those tests based on a larger number of lags. Using the more robust panel unit root test of Im et al. (2003), I additionally fail to reject the null hypothesis that all panels contain a unit root.

Figure 1.13: Domestic Rice Prices Follow a Unit Root Process

Tables

Table 1.25: Empirical Support for a Simple Model Explaining Spatial Variation in Rice Prices

	Dependent Variable: log rice price index$_t$ − log rice price index$_{t-k}$								
Growth Horizon	monthly ($k = 1$)			semi-annual ($k = 6$)			annual ($k = 12$)		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
log rice price index, $t - k$	-0.0293	-0.1042	-0.1601	-0.1827	-0.5795	-0.7984	-0.2532	-0.8065	-0.9877
	(0.0087)***	(0.0093)***	(0.0135)***	(0.0517)***	(0.0581)***	(0.0392)***	(0.0780)***	(0.0731)***	(0.0488)***
log shipping distance to THA/VNM via nearest port	0.0092	0.0119	-0.0001	0.0555	0.0527	0.0490	0.0867	0.0669	0.0858
	(0.0027)***	(0.0054)**	(0.0120)	(0.0149)***	(0.0247)**	(0.0675)	(0.0264)***	(0.0321)**	(0.0872)
post-2004m1	0.0671	0.2285	0.3767	0.5651	1.2520	1.4416	0.7914	1.5710	1.4745
	(0.0368)*	(0.0896)**	(0.1038)***	(0.1868)***	(0.3767)***	(0.2199)***	(0.2943)**	(0.4630)***	(0.2028)***
post-2004m1 × log shipping distance	-0.0089	-0.0222	-0.0356	-0.0545	-0.1068	-0.1127	-0.0810	-0.1301	-0.1026
	(0.0042)**	(0.0107)**	(0.0125)***	(0.0221)**	(0.0450)**	(0.0270)***	(0.0353)**	(0.0562)**	(0.0253)***
Wet Season Indicator	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
log Rainfall, $t-3 \ldots t-12$	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
City-Specific Linear Time Trend	No	Yes	Yes	No	Yes	Yes	No	Yes	Yes
City-Specific Quadratic Time Trend	No	No	Yes	No	No	Yes	No	No	Yes
City-Year Observations	2,638	2,638	2,638	2,638	2,638	2,638	2,638	2,638	2,638
R^2	0.502	0.526	0.547	0.639	0.732	0.792	0.691	0.812	0.868

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; Standard errors are clustered by city. The rice price index is produced by the Central Bureau of Statistics and was shared by Wimanda (2009). The wet season indicators vary across provinces as in Maccini and Yang (2009). Rainfall is calculated as the nearest observed monthly precipitation level. See notes to Figure 1.11 for details on the distance variable.

Table 1.26: Unit Values Confounded by Variation in Rice Subsidy Benefits

Dependent Variable:	average log unit value of rice in			
	2008		2005	
	(1)	(2)	(3)	(4)
share of households receiving Raskin, 2008	-0.284	-0.146		
	(0.054)***	(0.057)**		
share of households receiving Raskin, 2005			-0.203	-0.145
			(0.054)***	(0.068)**
constant	8.616	8.547	8.108	8.086
	(0.030)***	(0.029)***	(0.019)***	(0.021)***
Province FE	No	Yes	No	Yes
Number of Districts	416	416	379	379
R^2 (overall)	0.070	0.548	0.052	0.370

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; These estimates are based on a regression of the log unit value (in Rupiah) of rice in district j on the share of households in that district receiving rice subsidy (Raskin) benefits in the three months prior to survey enumeration in 2005 or 2008. Standard errors are clustered at the province level. Unit values could not be estimated in 2002 for districts in Aceh, the Riau Islands, Maluku, and North Maluku. For villages in these provinces, I therefore take the nearest province unit value observation in 2002.

1.9.6 Pareto Distributed Landholdings: Theory and Estimation

In this section, I provide additional background on the assumed Pareto distribution of landholdings as well as details on the empirical content of the estimated Pareto shape parameters $\widehat{\lambda}_v$. In the paper, I briefly discuss the empirical evidence pointing towards the Pareto distribution at the district level. A more systematic analysis of Paretian properties at the village level requires estimating distributional parameters for every village in the Agricultural Census.

I obtain estimates of the Pareto shape parameters, λ_v, for every village in Indonesia using the Gabaix and Ibragimov (2011) estimator. That is, for each village I regress the log rank minus 1/2 on the the log of the given landholding size. Given that some households within each village report the same landholding size, ties are broken by taking the average rank.[88] Identical results obtain when using the log minimum, log maximum rank, or the log complementary CDF as the dependent variable (the measures have mutual correlations above 0.95). In terms of differences in $\widehat{\lambda}_v$ across the three different measures of landholdings, Figure 1.14 demonstrates that total agricultural landholdings tend to yield the lowest estimates of λ_v (highest mean and dispersion) whereas wetland holdings tend to yield the largest estimates of λ_v (lowest mean and dispersion).[89] This is consistent with the existence of relatively smallholder rice agriculture and much larger plots used to grow other crops besides rice throughout the country.

Applying a test for departures from Paretian linearity suggested by Gabaix

[88]The discrete clumping at certain round landholding sizes is due in part to imperfect knowledge about plot sizes or boundaries. In practice, the assumed continuity of the Pareto distribution is a reasonable and mostly innocuous approximation to the discrete landholdings distribution—an assumption common in empirical work using the Pareto distribution (see Gabaix, 2009).

[89]In each case, there are a number of villages with $\widehat{\lambda}_v < 1$, which implies infinite mean landholdings under the strict Pareto assumptions. Although $\widehat{\lambda}_v \to \lambda_v$ under the usual OLS assumptions derived in Gabaix and Ibragimov (2011), sampling error explains most of the observed $\widehat{\lambda}_v$ below unity. When estimating λ_v using total agricultural landholdings, for example, nearly 4,000 villages have estimates of $\lambda_v < 1$. In all but 428 of these villages, however, the 95% upper confidence interval exceeds unity according to the unbiased standard error formula $\widehat{\lambda}_v \sqrt{2/N_v}$ given in Gabaix and Ibragimov (2011), where N_v is village population size.

(2009), I find that the strict Pareto assumptions do not hold in around 35 percent of villages. Indications of these departures can also be seen in Figure 1.15, which compares the distribution of $\widehat{\lambda}_v$ (for total agricultural landholdings) across villages for several different choices of $\underline{R}$. Results in Eeckhout (2004) suggest that the type of sensitivity to the effective choice of $\underline{R}$ evident in these figures potentially points to the underlying distribution of landholdings across *all* households (i.e., not simply those above $\underline{R}$) being log-normal rather than Pareto. Nevertheless, for reasons discussed in the paper, the Pareto provides a reasonable approximation to the landholdings distribution for the analytic purposes in this study. The goal is not to establish that landholdings irrefutably follow a power law in all villages, but rather that the formulation here provides a good fit to the data. Moreover, in Appendix 1.9.4, I showed that the key parameter estimates in the two-step model for flow migration rates are unaffected by imposing alternative choices of $\underline{R}$ in the estimation of λ_v.

Additionally, the estimated λ_v contain information on the distribution of wealth and agricultural activities in rural areas. Figure 1.16 shows the strong correlation between $\widehat{\lambda}_v$ and an expenditure-based measure $\widehat{G}_v$ of the income Gini coefficient *estimated without any information on landholdings.*[90] Figure 1.17 establishes further that λ_v is informative about the share of households engaged in the sale of agricultural products. If nearly all households in village v are subsistence consumers, then it is unlikely that changes in farmgate prices at the national or regional level will affect households residing there. Figure 1.17 confirms that villages with lower $\widehat{\lambda}_v$ (i.e., higher mean and greater variance in R_{iv}) are more likely to have a majority of households selling agricultural output. Figure 1.18 demon-

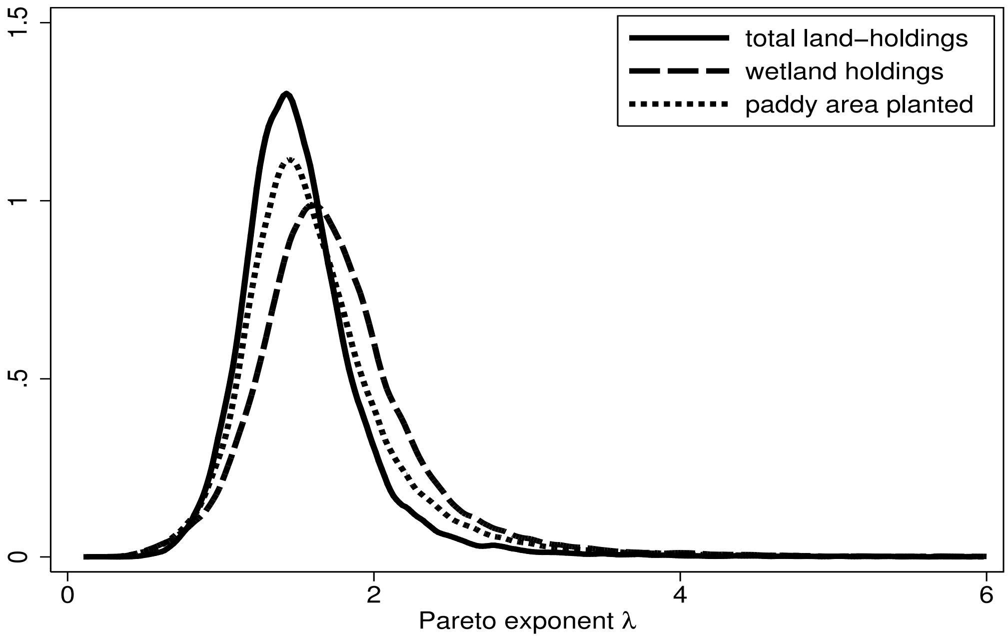
[90]Recall that the landholdings Gini for village v is given by $G_v(R_{iv}) = 1/(2\lambda_v - 1)$. The $\widehat{G}_v$ are obtained through the familiar Poverty Map procedure of predicting household expenditures per capita using observable variables in the Census. These estimates are the only available proxies for income inequality across all villages of Indonesia (see Suryahadi et al., 2005). The qualitative similarity of the nonparametric function with the exact mathematical expression for G_v in terms of λ_v can be partly explained by the (log-)linear relationship between landholdings and household expenditures/capita one finds in *Susenas* data. Interestingly, the implied Gini coefficient for landholdings based on $\widehat{\lambda}$ estimated across all Indonesian households is nearly identical to the Gini coefficient for household expenditures per capita estimated using *Susenas* for the last decade.

strates the nonlinear relationship between $\widehat{\lambda}_v$ and the extent of paid agricultural labor market activity. The shape of the relationship is quite intuitive. We find that the density of the paid agricultural labor market is (i) lowest in villages with the lowest mean and least inequality in landholdings (high $\widehat{\lambda}_v$), and (ii) highest in villages with intermediate dispersion in landholdings ($\widehat{\lambda}_v \approx 1.5$).

In developing the extensive margin results in the paper, I make use of the actual log maximum and minimum landholdings in each village, $\ln \widetilde{\underset{\sim}{R}}_v \equiv \ln \widetilde{R}_v - \ln \underset{\sim}{R}_v$. In Table 1.27, I demonstrate several properties of $\ln \widetilde{\underset{\sim}{R}}_v$ consistent with the Pareto distribution and empirical intuition more generally. The results suggest that $\ln \widetilde{\underset{\sim}{R}}_v$ is (i) increasing in the log population of the village, (ii) decreasing in $\widehat{\lambda}_v$, and (iii) increasing in the share of households with landholdings above $\underline{R} = 0.1$ Ha. These findings are consistent with the theoretical properties of landholdings extrema as discussed in the paper.

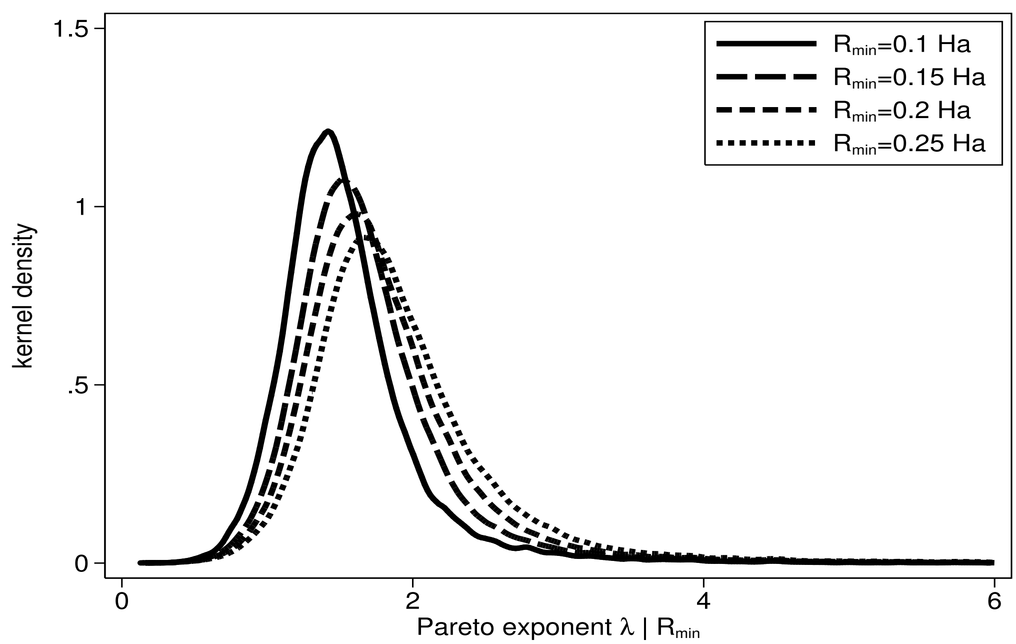
In closing, I mention a few stylized facts supporting the important assumption in the paper that the empirical landholdings distribution in village v is predetermined with respect to migration in v. First, note that the Agricultural Census was enumerated in late 2003 (i) two years before we first observe migrant stocks in *Podes* 2005, and (ii) several months prior to the initial discussion and eventual implementation of the import ban. Thus, the observed heterogeneity in landholdings could not be due to land transactions in expectation of or response to the price shock. Moreover, land markets in rural Indonesia are extremely thin. Benjamin (1995) demonstrates that farm sizes in Javanese villages are relatively fixed in the short-run due to imperfect land markets and long rental contracts. More recent *Susenas* data from 2005 covering the entire country confirm that less than one percent of households engage in land transactions over a one year horizon. The same transaction rate holds in the data from one year prior, suggesting that households had not purchased land in expectation of rising prices.

Figures



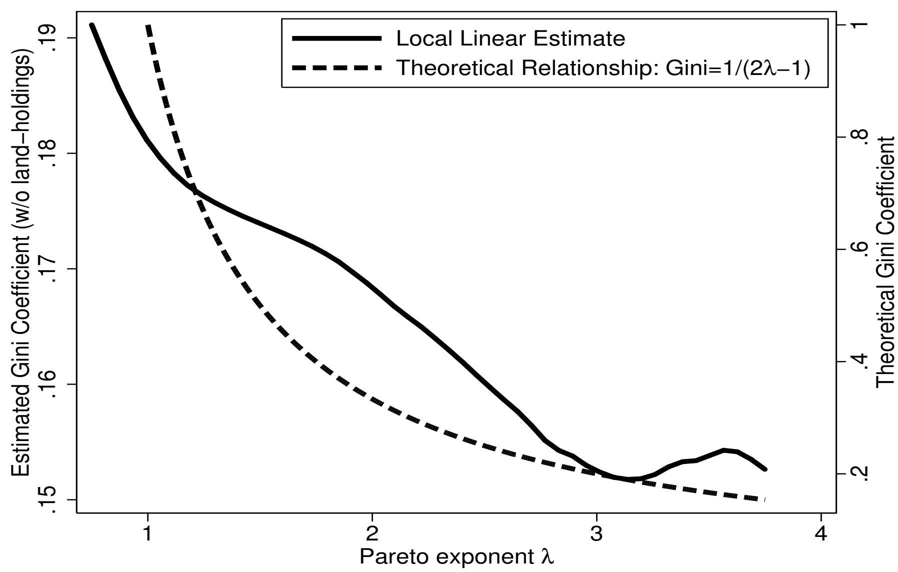
Notes: The figure plots kernel densities for $\widehat{\lambda}_v$ estimated with $\underline{R}$ = 0.1 Ha. The densities are based on an Epanechnikov kernel and a rule-of-thumb bandwidth.

Figure 1.14: The Distribution of $\widehat{\lambda}_v$ for Different Landholdings Measures



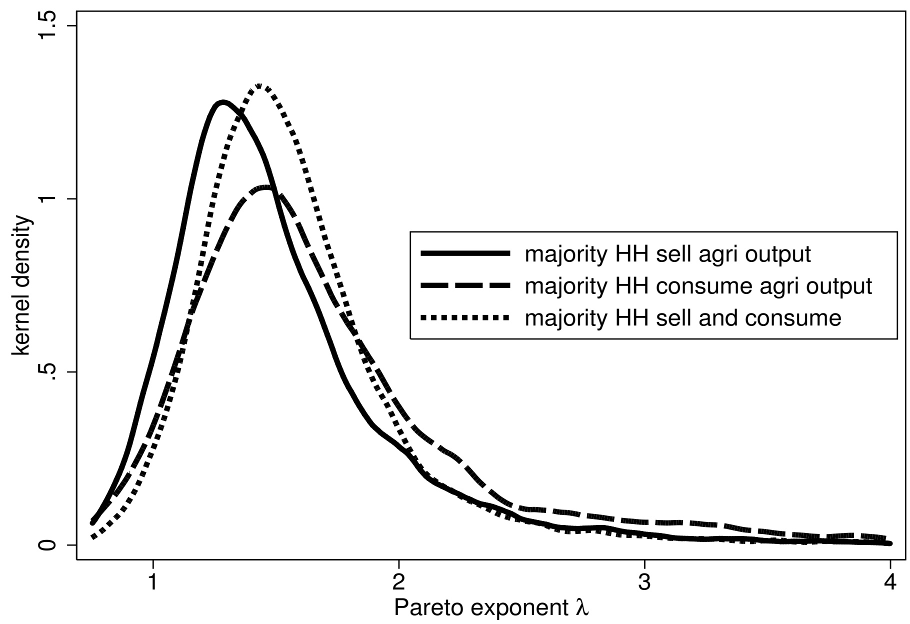
Notes: The figure plots kernel densities of the estimated λ_v (for total agricultural landholdings) obtained using several different choices of $\underline{R}$ and for comparison purposes restricting to those villages with non-missing estimates of λ_v for all $\underline{R}$.

Figure 1.15: Sensitivity of the Pareto Exponent to Lower Truncation $\underline{R}$



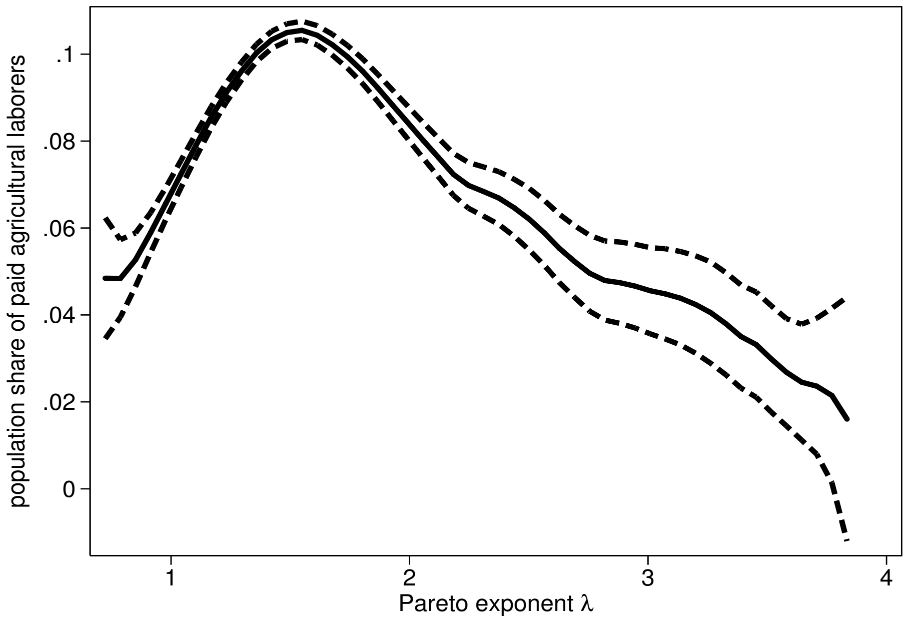
Notes: The solid line is the local linear estimate of the relationship between the estimated Gini coefficient $\widehat{G}_v$ and the estimated Pareto exponent $\widehat{\lambda}_v$ where $\widehat{G}_v$ is based on Poverty Map analytics and does *not* incorporate any information on household landholdings. See Suryahadi et al. (2005) for details. The local linear estimate employs a bandwidth of 0.15 and an Epanechnikov kernel while trimming the top and bottom 1% of $\widehat{\lambda}_v$. The dashed line plots the exact theoretical relationship between the Pareto exponent λ_v and the Gini coefficient G.

Figure 1.16: $\widehat{\lambda}_v$ and Non-Land-Based Measure of Income Inequality



Notes: The figures report nonparametric estimates of the probability that the head of village v reports in *Podes* 2005 that a majority of households in the village sell (net producers) or subsist (net consumers) on their agricultural production conditional on that village reporting agriculture being the most prominent source of employment. The estimated probabilities are based on local linear regressions using an Epanechnikov kernel, a bandwidth of 0.3, and a trimming of the top and bottom 1 percent of $\widehat{\lambda}_v$.

Figure 1.17: Probability Majority of Households in Village v Are Net Producers



Notes: The figure plots the nonparametric local linear regression of the share of the population engaged in paid agricultural labor against the estimated Pareto exponent $\widehat{\lambda}_v$ (for all agricultural land). The regression curve and 99% confidence band are based on an Epanechnikov kernel, a bandwidth of 0.125, and a trimming of the top and bottom 1 percent of $\widehat{\lambda}_v$.

Figure 1.18: $\widehat{\lambda}_v$ and the Extent of Paid Agricultural Labor Market Activity

Tables

Table 1.27: Log Difference in Max and Min Landholdings

Regional Fixed Effects	—	island	province	district	subdistrict
	(1)	(2)	(3)	(4)	(5)
log population of village v, 2005	0.468	0.510	0.491	0.482	0.493
	(0.020)***	(0.023)***	(0.022)***	(0.014)***	(0.012)***
Pareto exponent $\widehat{\lambda}_v$	-0.902	-0.881	-0.862	-0.798	-0.762
	(0.061)***	(0.059)***	(0.059)***	(0.057)***	(0.061)***
share households above $\underline{R}$	1.148	1.099	1.092	0.984	0.879
	(0.060)***	(0.059)***	(0.059)***	(0.051)***	(0.053)***
N	51,598	51,598	51,598	51,598	51,598
R^2	0.423	0.432	0.456	0.532	0.630

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%;   The table reports estimates from a regression of the form
$\ln(\widetilde{R}_v/\underset{\sim}{R}_v) = \alpha + \beta_1 \ln pop_v + \beta_2 \ln \widehat{\lambda}_v + \beta_3 share\ households\ above\ 0.1\ Ha_v + \kappa_j(v) + e_v$,
where $\kappa_j(v)$ is the given geographic fixed effect and also the level at which I cluster the standard errors.

1.9.7 Zeros, Balls, Bins, and Traveling Salesman

This section provides further background on the extensive margin, elaborating on the discussion of the balls-and-bins test and the stylized model of recruiters in the paper.

Ruling out a Balls-and-Bins Interpretation of the Extensive Margin

Although village population size proves to be one of the strongest determinants of the extensive margin, this relationship becomes less important for (i) villages exceeding some sufficiently large population threshold, and (ii) geographic units at higher levels of administrative aggregation. On (i), the sharp positive covariation between population size and the extensive margin flattens out when moving beyond 75th percentile of Indonesian village sizes ($N_v \approx 4,000$) (see Figure 1.19 below). Second, the share of administrative units with zero migration falls as one moves from the village to the level of the subdistrict or district, which comprise much larger populations than individual villages.[91]

Nevertheless, the extensive margin cannot be explained as a purely random phenomenon arising from the existing distribution of village sizes. To demonstrate this point, I adapt a simple probabilistic balls-and-bins test developed in Armenter and Koren (2010). The basic idea is to compare the empirical incidence of zeros with that arising from a model in which villages receive migrants (balls) randomly but with the probability proportional to village population size. Suppose that each migrant m is a ball. There are $M \in \mathbb{N}$ total migrants comprised of the sum across all villages, $M = \sum_{v=1}^{V} m_v$. Also, suppose that each village is a bin, the width of which is given by the share of that village's population in the total population of Indonesia. Formally, the size of bin v is given by $s_v = N_v/N$, where N_v is village v population and $N = \sum_{v=1}^{V} N_v$. The joint probability of migrants across villages

[91] The share of units reporting zero migrants falls from 46 (41) when examining villages to 16 (13) percent of 5,112 subdistricts reporting no migrants in 2005 (2008). At the district level, the number of zeros fall to 6 (3) percent of 409 districts in 2005 (2008).

follows a multinomial distribution

$$\mathbb{P}(m_1, \ldots, m_V) = \frac{M!}{m_1! \cdots m_V!} s_1^{m_1} \cdots s_V^{m_V},$$

in which the expected number of *nonzero* migration villages $\mathcal{V}^*$ (or non-empty bins) is given by

$$\mathbb{E}[\mathcal{V}^*|M] = \sum_{v=1}^{V} \left[1 - (1 - s_v)^M\right]. \tag{1.30}$$

Calculating the sample analogue of (1.30) for 2008, I find that the balls-and-bins model predicts nonzero migration in 64,457 villages out of the total 65,966 villages. (If population sizes were uniform across villages, the balls-and-bins model predicts that every village would almost surely have at least one contract migrant.) In other words, only 5.5 percent of the 27,297 zero migration village in the empirical data can be explained away as an atheoretical statistical regularity in sparse data.[92] Figure 1.20 compares the predicted probability of having any migrants under the balls-and-bins model (dashed curve) with the actual share of villages with any migrants (solid curve). Both are plotted against log village population size. The incidence of zeros in the data is much higher than would be predicted on the basis of a random balls-and-bins allocation of migrants across villages. The vertical distance between the two curves constitutes the scope for the theory and empirics in the paper to address the substantive economic forces behind the extensive margin including, among others, the role of recruiters.

A Heuristic Framework for Recruiter Location Choice

To illustrate the logic behind the traveling salesman implications discussed in Section 4.1 in the paper, first consider two districts k and k' with equal pop-

[92] Performing separate balls-and-bins calculations within each province and then aggregating across provinces reveals that 13 percent of villages have zero migrants or that 33 percent of empirical zeros can be deemed statistical artifacts. These figures increase to 20 and 50 percent, respectively, when looking within districts. So, the level at which migrants are randomly assigned matters. Regardless, the incidence of atheoretical zeros compares favorably with the trade data considered by Armenter and Koren (2010) in which nearly 90 percent of the zeros found in U.S. destination×10 digit product level trade data could be explained using a balls-and-bins framework.

ulations and inter-village travel distances. District k has two equally populated villages, while district k' has three villages: village $1_{k'}$ has equal population with the two villages in k, while villages $2_{k'}$ and $3_{k'}$ are equally populated with the total equal to the population of $1_{k'}$. Assuming (i) constant fixed costs of establishing agency presence in equally sized districts and (ii) constant fixed costs of entering villages, a given recruiter would be more likely to enter district k than k'. If, however, recruiters choose to visit district k' for other (unobserved) reasons, then village $1_{k'}$ would be more likely visited than $2_{k'}$ or $3_{k'}$. Now, add one identical village to each district with population greater than all existing villages in each district. Assuming recruitment agencies are subject to budget constraints preventing visits to all villages within k, it is straightforward to show within this framework that recruiters would only visit the newly added village in k.

I now provide a formal sketch of the general traveling salesman model underlying the proposed instruments discussed in Section 4.1. Begin by considering the problem of recruiters selecting a district within which to operate, retaining the assumption that licensing and other fees are paid to district government officials.[93] Let the cost of traveling between villages v and v' within district k be denoted by $d^k_{vv'} > 0$. Suppose further that there are $\mathcal{V}^k$ villages in district k and that the population of the district less village v is given by N^k_{-v}. For empirical tractability, additionally assume that the least-cost path of visiting every village within a district (in the sense of solving the NP-hard traveling salesman problem) is approximately proportional to the area of the district less village v, A^k_{-v}.[94] Now suppose that the fixed cost of entering a village, f, is identical across all villages. Suppose also that the ex ante probability of successfully recruiting any individual

[93]This assumption follows from evidence on the procedures through which recruitment agencies engage with government institutions in the process of authorizing legal contract emigrants (Bank Indonesia, 2009). The engagement with local government officials has been increasing in recent years as decentralization has resulted in a devolution of authorities and regulatory power to the regions (see Bachtiar, 2011). Under Law 39/2004, recruitment agencies are only permitted to recruit and place prospective labor migrants who are registered at the local Ministry of Manpower and Transmigration. Of course agency field workers often bypass local governments and enter villages directly, but the agencies must still liaise with officials in the district capital for the purposes of document preparation and other predeparture certification processes.

[94]That is, $\sum_{v'\neq v}^{\mathcal{V}^k} \sum_{v\neq v'}^{\mathcal{V}^k} d^k_{vv'} \propto A^k_{-v}$.

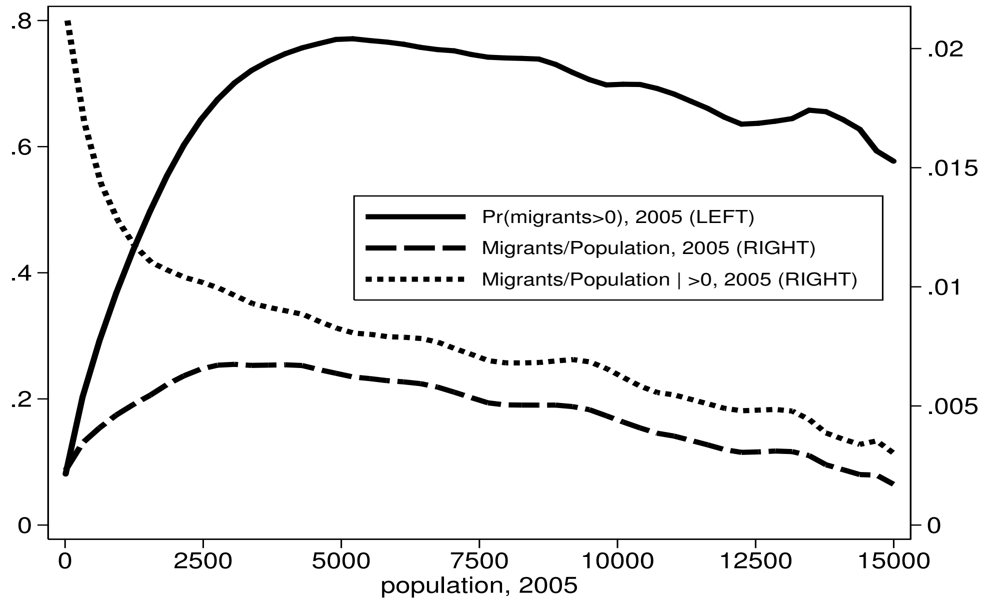
potentials migrant is identical across villages. The optimization problem for the recruitment agency is to maximize the number of potential migrants, $\widetilde{M}_v$, reached (with advertisements and contract offers) and to minimize the costs (and hence maximize expected revenue). The objective is then

$$\max_k \sum_{v=1}^{\mathcal{V}^k} \widetilde{M}_v \quad \text{s.t. } f\mathcal{V}^k + A^k_{-v} \leq \mathcal{B}, \quad \widetilde{M}_v \leq N_v \forall v$$

where $\mathcal{B}$ is the exogenously given budget of the recruitment agency. The (heuristic) solution function (i.e., the optimal district) should be increasing in N^k_{-v} and decreasing in A^k_{-v} and $\mathcal{V}^k$. Once inside a given district k, all else equal, budget-constrained recruiters are relatively more likely to visit villages with larger populations since the unconditional probability of successfully recruiting a single migrant is higher.

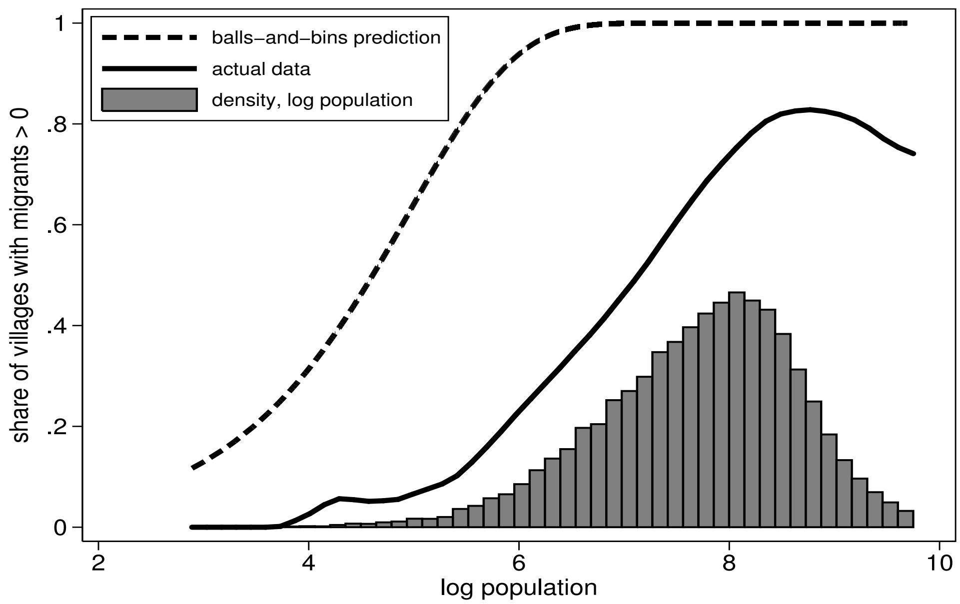
In Table 1.28, I test the stylized traveling salesman model for recruiter visits using the only available proxy for such visits—a measure in Village Potential data form 2008 indicating whether any recruiters targeting female migrants visited the village in the prior year. Conditional on the presence of any migrants in the prior period (i.e., as recorded mid-2005), the likelihood that village v is visited by a recruiter in period $t+1$ (i.e., prior to mid-2008) is (i) increasing in the population of village v, the population of v's district less v, and (ii) decreasing in the the number of village in v's district. These findings are consistent with the predictions of the traveling salesman framework sketched above. The positive albeit statistically null correlation with the area of v's district is not. However, that the probability is decreasing in the distance from the subdistrict capital (albeit not the district capital) suggests that some of the distance components of the model hold. A more rigorous test would require computing the actual distances between villages and using some of the available methods for solving the traveling salesman. This is beyond the scope of the present study as the patterns observed in Table 1.28 bear out indications that recruiter decisions follow some approximation to the model described above. Of course, the unobserved determinants of recruiter choice likely swamp those included below. Nevertheless, other features worth noting are (i) the positive correlation with the share of ethnic Arabs in the total population of v in 2000, (ii) the negative correlation with the share of post-primary educated, and (iii) the statistically null correlation with agricultural income shocks. Findings (i) and (ii) are consistent with the low skill levels and high propensity towards Middle Eastern destinations among Indonesian migrants. The latter finding (iii) suggests that recruiters do not necessarily respond to income shocks in deciding where to target their contract offers.

Figures



Notes: The estimates are based on local linear regressions using an Epanechnikov kernel, bandwidth of 400, and trimming the top 4 percent of villages for presentational purposes. The figure on the right looks similar when expressing either variable in logs. The 95% confidence bands are calculated using analytic standard errors.

Figure 1.19: Population Thresholds in the Margins of Migration



Notes: The "actual data" estimate is based on a local linear probability regression of an indicator for any migrants in village v on log population, using an Epanechnikov kernel and trimming the top 1 percent of villages for presentational purposes. The "balls-and-bins" prediction is based on the model described in Section 1.9.7.

Figure 1.20: Actual Incidence of Zeros vs. the Balls-and-Bins Prediction

Tables

Table 1.28: On the Determinants of Recruiter Visits

$\mathbb{P}$(female migrant recruiter visit by $t+1$)		
	(1)	(2)
$\mathbf{1}$(migrants in $t > 0$)	0.116	0.109
	(0.011)***	(0.009)***
log village population, t	0.048	0.046
	(0.004)***	(0.004)***
log district population less v, t	0.041	0.047
	(0.013)***	(0.015)***
log number of villages in district	-0.018	-0.026
	(0.014)	(0.015)*
log district area less v	0.011	0.011
	(0.008)	(0.007)
Pareto exponent $\hat{\lambda}_v$		-0.001
		(0.005)
share households above $\underline{R}$		-0.022
		(0.018)
log distance to subdistrict capital		-0.012
		(0.003)***
log distance to district capital		-0.004
		(0.006)
Arab population share		1.023
		(0.464)**
Chinese population share		0.697
		(2.221)
Muslim population share		-0.041
		(0.030)
share with post-primary education		-0.108
		(0.043)**
price shock, t		-0.212
		(0.174)
rainfall shock, t		-0.018
		(0.014)
Number of Villages	51,593	51,593

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The table reports marginal effects at the mean from a probit regression of all covariates shown in the respective columns as well as the following additional covariates in column 2: wetland area as a share of total farmland, log distance to nearest emigration center, and an indicator for government-prescribed urban status. Standard errors are clustered at the district level.

1.9.8 Panel Data Construction

In this subsection, I describe the process of constructing a panel dataset of Indonesian villages comprised of data collected in 2000, 2002, 2003, 2005, and 2008. Starting from the baseline 65,966 villages in Table 1 in the paper, the final sample of villages is reduced further by two factors. First, because this paper focuses on heterogeneous income shocks in agricultural areas, I exclude urban villages without land-holdings entries in the Agricultural Census. There are other practical reasons for doing so as well. In Indonesia, agricultural commodity price increases generally have homogeneous, negative effects on real income in urban areas, and (ii) rainfall shocks tend to have null effects on rice production in nominally urban areas (Levine and Yang, 2006).

Second, changes in administrative boundaries over the period 2000-2008 required dropping a small number of villages with missing data from one or more of the additional sources, including the Population Census of 2000. I ultimately treat these villages as missing at random. In the late 1990s and early 2000s, responding to a range of political and economic incentives in the wake of decentralization, government officials set about proliferating administrative units across the country and at varying levels of government (see Fitrani et al., 2005). The proliferation was relatively more common in the Outer Islands than on Java. This process has created difficulties for researchers attempting to link administrative units over time in the *Podes* and other surveys. Most researchers work with district-level aggregates and take districts in some base year and aggregate backwards to achieve minimum comparative areas (MCA) (e.g., Vothknecht and Sumarto, 2009). For studies such as the present one, however, it is crucial to retain the village as the unit of analysis.

The remainder of this appendix details the matching of villages across multiple waves of *Podes* (2002, 2005 and 2008), the 2003 Agricultural Census, and the 2000 Population Census. Prior to beginning, I exclude villages from the islands of Papua in Eastern Indonesia and Nias off the West coast of Sumatra. I exclude Papua because the data quality is questionable and moreover the social and economic conditions do not lend themselves to the issues addressed in this study. I exclude Nias since the special post-Indian Ocean tsunami *Podes* survey administered in this region in mid-2005 did not include questions on migration.

Panel construction proceeds with the merging of villages recorded in *Podes* 2005 and 2008. The Central Statistics Bureua (BPS) does not provide exact concordances between villages across these two survey rounds. As such, I manually construct a mapping between these two data sources, which contain the main dependent variables of interest, using a combination of exact and fuzzy merge-matching algorithms combining information on province ID, district ID, subdistrict ID, village ID, village name, and village land area. In the initial step, I combine 2008 villages with identical 2007 village IDs as made available in *Podes* 2008. I also remove all non-diacritic characters from village names prior to implementing the algorithm. The resulting panel is comprised of 65,966 MCAs. A detailed breakdown by province of the number of villages matched at each stage of the algorithm can be seen in Table 1.29. Around 700 villages could not be merged into a reliable MCA across years.[95] Nevertheless, I view these villages as missing at random insomuch as the timing of elections resulting in the splitting of districts and subsequently villages has been shown elsewhere to be orthogonal to baseline observables of interest (Skoufias et al., 2010).

At the next stage of matching, I incorporate data from the 2000 Population Census using the unique administrative IDs available in *Podes* 2005. The merge-matching algorithm proceeds analogously to that described above. Given the relatively longer period of possible administrative proliferation between 2000 and 2005, the resulting success rate in matching villages was lower than that obatined for *Podes* 2005 and 2008. I then repeat the matching procedure for villages recorded in *Podes* 2002, which contains the requisite information to construct the commodity price index (sans rice) used in Appendix 1.9.4. The resulting match rate is again less favorable than that obtained for *Podes* 2005 and 2008. Table 1.30 shows the

[95] These villages account for 4,570 migrants in 2005 and 12,746 migrants in 2008.

match rates by source and province.

Lastly, I perform a similar matching procedure for the Agricultural Census conducted in August 2003. In rural areas, all agricultural households were enumerated in every village. In urban areas as defined by the government, households in a sample of villages among those with any agricultural activities received enumerators. Additionally, due to security concerns at the time, only a subsample of all households in a few villages were enumerated in Aceh Province in May 2004. Due to a lack of village names for certain areas in a few provinces in the raw data provided by BPS, I was unable to merge a number of villages in the Agricultural Census using the *Podes* 2005 IDs (or in unreported results, the *Podes* 2002 IDs). Ultimately, however, the resulting panel consisting of data from 2000, 2002, 2003, 2005, 2008 comprises the overwhelming majority of Indonesian villages and particularly those in rural areas where international migration constitutes an important labor market opportunity.

Tables

Table 1.29: Merge-Matching Procedure for Linking *Podes* 2005 and 2008

	Matching Stage								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	Total
Matching Variable	*Type of Matching*								
Province ID	Exact	Exact	Exact	Exact	Exact	Exact	Exact	Exact	
Kabupaten ID	Exact	Exact	Exact	Exact	Exact		Exact		
Kecamatan ID	Exact	Exact	Exact	Exact					
Village ID	Exact								
Village Name	Exact	Fuzzy	Exact	Fuzzy	Exact	Exact	Fuzzy	Fuzzy	
Land Area		Exact							
Aceh	5047	153	64	99	161	174	128	109	5935
North Sumatra	3216	223	65	162	161	588	250	179	4844
West Sumatra	522	165	79	79	2	0	30	3	880
Riau	1183	20	116	34	29	0	81	4	1467
Jambi	871	25	57	39	66	0	151	5	1214
South Sumatra	1553	21	181	46	202	34	521	104	2662
Bengkulu	813	9	18	16	61	1	281	6	1205
Lampung	1661	37	33	58	68	84	188	45	2174
Kepulauan Bangka Belitung	305	6	1	8	0	0	0	0	320
Kepulauan Riau	170	2	5	8	11	0	49	0	245
DKI Jakarta	264	0	0	3	0	0	0	0	267
West Java	5036	22	392	53	43	52	91	103	5792
Central Java	8442	7	18	49	16	0	28	1	8561
Yogyakarta	438	0	0	0	0	0	0	0	438
East Java	8302	35	13	75	11	0	38	1	8475
Banten	1154	2	77	10	38	39	129	28	1477
Bali	681	2	2	5	8	0	3	0	701
West Nusa Tenggara	688	11	26	39	8	0	45	0	817
East Nusa Tenggara	1828	69	123	47	176	181	129	160	2713
West Kalimantan	1258	24	13	22	21	57	26	72	1493
Central Kalimantan	1087	16	22	23	63	0	106	8	1325
South Kalimantan	1634	27	122	11	74	0	75	3	1946
East Kalimantan	978	23	146	46	28	1	74	16	1312
North Sulawesi	942	2	8	11	22	97	52	121	1255
Central Sulawesi	1308	19	0	24	102	0	67	3	1523
South Sulawesi	2257	196	34	113	80	156	79	282	3197
Sulawesi Tenggara	1089	13	43	27	70	65	248	101	1656
Gorontalo	351	0	0	10	5	8	23	49	446
Maluku	621	42	80	54	21	3	17	27	865
North Maluku	533	2	16	7	48	0	144	11	761
Total	54232	1173	1754	1178	1595	1540	3053	1441	65966

Notes: This table reports the number of villages matched at each stage of the algorithm I devised in order to link *Podes* 2005 and 2008. Fuzzy matching was done using the `reclink` program with a minimum match score of 0.6 followed by visual inspection and manual matching at each stage of the process.

Table 1.30: Percentage of Villages in 2005 with Missing Data by Source and Province

Province (2005)	Population, 2005	# Villages, 2005	Agri. Census, 2003	Podes, 2002	Census, 2000	Pov. Map, 2000
Aceh	4,115,642	5,935	0.986	0.062	0.645	0.681
North Sumatra	11,258,698	4,844	0.104	0.151	0.057	0.069
West Sumatra	4,464,086	880	0.374	0.094	0.256	0.275
Riau	4,647,206	1,467	0.076	0.087	0.166	0.174
Jambi	2,675,052	1,214	0.056	0.050	0.081	0.081
South Sumatra	6,684,582	2,662	0.099	0.231	0.092	0.102
Bengkulu	1,579,034	1,205	0.285	0.469	0.075	0.083
Lampung	7,184,793	2,174	0.024	0.032	0.068	0.070
Kepulauan Bangka Belitung	1,023,689	320	0.459	0.547	0.116	0.131
Kepulauan Riau	1,082,151	245	1.000	1.000	0.335	0.331
DKI Jakarta	7,484,573	267	0.004	0.000	0.007	0.007
West Java	37,355,255	5,792	0.010	0.013	0.014	0.018
Central Java	32,771,370	8,561	0.012	0.002	0.004	0.007
Yogyakarta	3,407,430	438	0.014	0.000	0.000	0.000
East Java	35,907,891	8,475	0.015	0.002	0.004	0.013
Banten	8,809,337	1,477	0.006	0.003	0.005	0.005
Bali	3,271,583	701	0.021	0.021	0.034	0.036
West Nusa Tenggara	4,227,864	817	0.716	0.140	0.147	0.214
East Nusa Tenggara	4,265,106	2,713	0.095	0.118	0.091	0.164
West Kalimantan	3,973,249	1,493	0.132	0.169	0.073	0.128
Central Kalimantan	1,897,154	1,325	0.026	0.063	0.081	0.279
South Kalimantan	3,203,372	1,946	0.032	0.145	0.011	0.079
East Kalimantan	2,870,860	1,312	0.072	0.079	0.212	0.252
North Sulawesi	2,168,461	1,255	0.255	0.325	0.116	0.122
Central Sulawesi	2,362,476	1,523	0.099	0.135	0.099	0.101
South Sulawesi	8,338,093	3,197	0.121	0.116	0.101	0.118
Sulawesi Tenggara	1,975,490	1,656	0.193	0.365	0.091	0.124
Gorontalo	907,503	446	0.368	0.413	0.184	0.188
Maluku	1,343,401	865	0.101	0.351	0.086	0.135
North Maluku	875,599	761	0.419	0.717	0.252	0.355

Notes: This table reports the rate of failed matches between villages reported in *Podes* 2005 and the given source of data listed in the top row of the table.

Chapter 1, in part is currently being prepared for submission for publication of the material, Bazzi, Samuel. The dissertation author was the primary investigator and author of this material.

Chapter 2

It's All in the Timing: Cash Transfers and Household Expenditures in a Developing Country

Abstract

We use a large-scale cash transfer program in Indonesia to demonstrate the importance of timing and expectations in evaluating fiscal interventions. Expenditures respond primarily to unexpected changes in transfer income: Timely receipt of transfers yields no expenditure change relative to non-recipients. However, delayed receipt of the second transfer yields a large reduction in expenditures by roughly 7.5 percentage points. We reconcile these results with a consumption smoothing model that can be generalized to other settings. We provide further empirical support for this framework and argue that the asymmetric response to positive and negative shocks is consistent with binding liquidity constraints.

2.1 Introduction

Cash transfer programs have been a popular policy tool in developing countries over the last decade. Beyond their role in fostering human capital investment, cash transfers are increasingly viewed also as a potential vehicle for stimulating or sustaining household expenditures, often in the process of introducing broader policy changes (e.g., the reform of regressive subsidies, see Coady et al., 2010). A large body of work evaluates transfer programs in the United States through the lens of the life-cycle/permanent income hypothesis (hereafter, PIH) and shows how the expenditure impacts depend on the timing and expected duration of transfers (see Jappelli and Pistaferri, 2010). Yet, we have limited evidence on whether the effectiveness of fiscal intervention in low-income countries hinges on similar factors implied by our canonical models of consumption behavior.

In this paper, we aim to fill this gap by asking how household expenditures evolve over the course of a cash transfer program with regularly scheduled disbursements in a developing country context. We investigate this question both theoretically and empirically in the context of a large-scale, unconditional cash transfer (UCT) program in Indonesia providing nearly 19 million households with quarterly transfers of around 30 USD for a period of one year in the wake of fuel subsidy cutbacks.[1] Several unique administrative features of the program make it possible to test for consumption smoothing behavior predicted by the PIH. However, our findings yield several interesting implications for the evaluation of programs beyond our specific setting.

We begin by developing a general evaluation framework based on testable implications of a simple version of the PIH with transitory income transfers. The key insight of the model is that in the absence of liquidity constraints, household expenditures should only respond to unexpected changes in transfers. This result gives rise to several interesting possibilities. First, it suggests that households may begin to change expenditures upon learning about a subsequent program even

[1] At roughly one-eighth of average household expenditures, UCT benefits were slightly smaller than the well-known *Progresa* transfers, which comprised roughly one-third of household income at baseline.

prior to the initial disbursement of cash. Second, unforeseen delays in disbursements may lead to negative expenditure shocks as households draw down savings ahead of the scheduled transfer date. Third, if transfers arrive on schedule, then beneficiaries with rational expectations should not exhibit differential expenditure growth relative to non-recipients.

Using this conceptual framework and nationally representative household-level panel data, we examine the expenditure impact of Indonesia's UCT over two time horizons: (i) a short-term period after which all beneficiary households had received at least one transfer and (ii) a medium-term period spanning several months after the program ended. Our identification strategy relies on several sources of exogenous variation in the incidence, timing, and scale of transfer income.

First, many non-poor households received the UCT while many poor did not, making it possible to construct a credible counterfactual using a difference-in-difference reweighting approach (Abadie, 2005).[2] Given pervasive targeting errors, reweighting non-recipient households by their estimated odds of treatment effectively rebalances treatment and control households along baseline expenditure levels. Conceptually, this ensures that recipient and non-recipient households draw randomly from the same *ex ante* expenditure distribution, which is essential for testing the theory.

Second, due to arbitrary administrative delays, the second quarterly transfer was staggered across regions.[3] Given variation in local disbursement schedules, nearly 30 percent of all recipients in our sample were still awaiting their second transfer at the time of enumeration in early 2006. We show that the staggering process cannot be explained by any observable differences across regions in terms of remoteness, weather shocks, level of development, or political affiliation.

Third, because all households received the same transfer amount per dis-

[2]Although beneficiaries were targeted through a quasi-means testing process, we show that the reconstructed proxy means test (PMT) scores used to assign eligibility are too weakly correlated with treatment to justify a fuzzy regression-discontinuity design. Our augmented model for predicting program receipt captures substantial variation in treatment status across households and, in fact, outperforms the approximate PMT scores.

[3]Indonesia's administrative divisions proceed from province to district to subdistrict to village. The staggering took place primarily across subdistricts, the level at which post offices were tasked with disbursing quarterly transfers to villages.

bursement, we observe considerable variation in transfers *per capita*. This allows us to identify an auxiliary measure of the marginal propensity to consume (MPC) out of transfer income. De Janvry and Sadoulet (2006) make use of similar variation in treatment intensity imposed by the cap on total transfers in the *Progresa* program in Mexico, and Kaboski and Townsend (2005, 2011) analogously exploit fixed financial transfers across Thai villages that vary in population size.

Our central empirical results suggest that timing and expectations matter. On the one hand, recipient households still awaiting their second quarterly transfer in early 2006 report per capita expenditure growth rates that are roughly 0.075 log points lower on average than both (reweighted) control households and UCT beneficiaries that had already received the second transfer. The drop in expenditures for a subset of beneficiaries can be interpreted as the effect of income falling short of forecasted expectations. In other words, savings were drawn down too early ahead of the (delayed) transfer. Although the income shock was transitory, households with limited borrowing options were then forced to adjust expenditures downward. On the other hand, we find no mean differences in expenditure growth between control households and UCT beneficiaries that had received the full two transfers as expected by early 2006. That is, timely receipt of UCT disbursements had no economically significant effect on expenditure changes. Overall, the expenditure impacts of a positive transitory shock associated with the initial rollout of the UCT program are much smaller than the impacts of the negative shock associated with delayed disbursement of the second quarterly transfer. This asymmetric response is consistent with consumption smoothing behavior in the presence of credit constraints.

Several other findings support the model-based interpretation of our main results. First, the largest differential growth rates through early 2006 are found for food rather than non-food expenditures. This is consistent with the cash-on-hand implications of the model since food expenditures are reported over the prior week whereas non-food expenditures are reported over the prior month (or year). Second, the negative expenditure impact of delayed disbursement is increasing over time with larger effects observed for households enumerated in March than in

February 2006. Third, the differential impacts observed midway through the program disappear by early 2007, nearly six months after the final quarterly transfer was received by all beneficiaries. Fourth, proximity to financial institutions largely offsets the negative shock associated with delayed disbursements. For beneficiaries residing in areas with more liquid local financial markets, the delays did not lead to any significant change in mean expenditures. Fifth, we find a U shaped relationship between the estimated treatment effects and household age—a pattern that is consistent with life cycle versions of the PIH in which the MPC is highest among the young and old. Finally, using the variation in transfers per capita, we estimate an auxiliary MPC out of transfer income of around 0.10, which is in line with the MPC estimated based on transfer incidence.

Moreover, our main findings are not an artifact of the data. All of the key null results for both the short- and medium-term are precisely estimated zeros, and we demonstrate that the sample sizes are sufficiently large to detect small treatment effects. We also show that spillovers to non-recipients via informal taxation and redistribution within villages cannot explain the results.[4] The baseline results hold up to a battery of robustness checks including the use of a measure of non-food expenditures over the past year rather than the past month. This partially rules out concerns that transfer funds were spent during a period missed by the survey instrument.

This paper finds several interesting parallels with the literature on fiscal interventions in high-income countries. We contribute to a growing literature that quasi-experimentally tests theories of consumption and offers new evidence on how timing and expectations matter for understanding policy impacts. Hsieh (2003), for example, shows that household expenditures in Alaska do not respond to regular cash transfers provided through the Alaska Permanent Fund but are excessively sensitive to occasional tax rebates. Browning and Collado (2001) find a similar non-response to regular semi-annual extra wage payments to full-time workers in

[4] It is also important to note that even if some poor households in the control group were expecting to receive the program, the model suggests that we should observe negative expenditure growth among these households thereby pushing down the control group average and working against our main findings.

Spain. Several recent studies exploit the randomized timing of cash transfers to American households. Johnson et al. (2006) find that on average households spent 20-40 percent of the tax rebates in 2001; Parker et al. (forthcoming) show that 12-30 percent of economic stimulus payments (ESP) in 2008 were spent primarily on nondurable expenditures. In both cases, the large response to these transitory income shocks occurred immediately after disbursement and was driven by households with low liquid wealth and income. These findings are broadly consistent with ours: (i) Indonesian households respond primarily to unexpected changes in (the timing of) transfer income, and (ii) the largest MPC is found for nondurable food expenditures, and (iii) liquid local financial markets facilitate *ex post* consumption smoothing against unanticipated negative income shocks. Moreover, as in the ESP studies, low-frequency data constraints make it difficult to identify causal expenditure impacts beyond those immediately correlated with the timing of disbursements.[5]

We also directly contribute to a small but growing body of work showing that explicitly accounting for timing and expectations can enrich our understanding of the mechanisms underlying the average treatment effects of cash transfer programs in developing countries. Bianchi and Bobba (forthcoming) show that *Progresa* increased entrepreneurial activity among Mexican beneficiaries in advance of actual transfer receipt. By exploiting the differential timing of transfers across households, they are able to show that the program increased entrepreneurship not only by relaxing liquidity constraints but also by encouraging risk-taking. Using a similar identification strategy, Edmonds (2006) finds that South African households reduce child labor and increase schooling in anticipation of future transfers from the Child Support Grant Programs, attributing the result to binding liquidity constraints. Beyond these reduced-form studies, Attanasio et al. (2012) develop a structural model that makes it possible to assess the effect of control households' expectations over future transfers on observed treatment effects. Failing to account for such expectations can lead researchers to underestimate actual treatment effects. Although this bias did not arise in the case of *Progresa*, the Attanasio et al.

[5]Thus, like Broda and Parker (2012), we emphasize the partial equilibrium nature of the analysis and refrain from making claims about changes in aggregate consumption.

framework resonates with our results showing that failure to account for unmet expectations over the timing of transfers lead one to underestimate the expenditure impacts of timely transfer receipt by 30 percent.

Like ours, each of these three studies suggests large potential research gains to (re)optimizing survey and study design. Recent work has explored optimal program design along several dimensions.[6] Alongside these important innovations, our results suggest that it may be possible to learn more from existing designs by modeling and testing behavior change that occurs outside typical enumeration windows.[7] In fact, our findings raise the possibility that existing studies may actually understate the full (expenditure) gains of cash transfers by focusing on impacts observed at a narrow point in time. Going back to the title of the paper, it is not only the case that program impacts may depend crucially on the timing of interventions but also that our understanding of those impacts as researchers may hinge on the time(s) at which we observe households.

For related reasons, our findings also bolster the case for alternative design of (non-)experimental impact evaluations in McKenzie (2012). When outcomes of interest have low autocorrelation (e.g., expenditures), McKenzie shows that more follow-up surveys conducted over relatively short intervals can dramatically increase the power to detect small treatment effects. Beyond these mechanical statistical gains, however, additional follow-up surveys can substantially increase the value added of a given impact evaluation. In our study, for example, higher frequency surveys would have allowed us to test additional implications of the theory. Overall, our paper contributes to a growing literature arguing for greater attention to pre-intervention (post-announcement) behavior change, differential treatment effects within the course of an intervention, and possible backloading of

[6]For example, Baird et al. (2011) evaluate the relative effectiveness of conditional versus unconditional cash transfersBarrera-Osorio et al. (2011) compare the effectiveness of cash transfers conditional on re-enrollment and graduation relative to a standard design conditional on attendance, Carrillo and Ponce Jarrín (2009) model and evaluate the relative effectiveness of alternative cash delivery mechanisms, and Filmer and Schady (2011) compare the impact on school attendance of conditional cash transfers that vary in size.

[7]That is, initial observations from a baseline conducted prior to announcing the program and follow-up observations from a single endline conducted after the given intervention has been active for some time if not already ended.

changes post-intervention via accumulated (or depleted) savings.

The remainder of the paper proceeds as follows. Section 2.2 develops a simple theoretical model of consumption smoothing with testable implications for the study of cash transfer programs. Section 2.3 provides background on the empirical setting and details our identification strategy. Section 2.4 presents the empirical results, and Section 2.5 concludes.

2.2 Cash Transfers and Consumption Smoothing: A Simple Model

In many contexts, cash transfers constitute a substantial change in household income. In this section, we develop a simple model of consumption smoothing in which the impact of transfers to expenditures depends on the timing of these transfers as well as household expectations. The baseline setup relies on a standard version of the permanent income hypothesis (PIH). In keeping with the empirical context in many evaluation settings, we consider expenditure growth between two (survey enumeration) periods and abstract away from permanent income shocks.[8] The timing convention is developed in a manner consistent with the UCT program we evaluate in this paper, but it is straightforward to generalize to other settings with alternative disbursement frequencies, delays, and forecasting horizons.

Under a standard set of assumptions, the model delivers the following equation[9]

$$\Delta \ln C_{ht} = \left(\frac{r}{1+r}\right) \underbrace{(\ln Y_{ht} - \mathbb{E}_{t-1} \ln Y_{ht})}_{\text{income shock}}, \tag{2.1}$$

which relates growth in household h's log consumption to the real interest rate (r) and the difference between realized and expected log income where $\ln Y_{ht} =$

[8]This is without loss of generality insomuch as recipients and non-recipients are *ex ante* identical in terms of long-run income prospects under the assumptions necessary to recover causal effects in a standard difference-in-difference setting (either experimental or non-experimental). See Section 2.3.3.

[9]We start from the usual Euler equation for consumption $u'(C_{h,t-1}) = (1+\delta)^{-1}\mathbb{E}_{t-1}\left[(1+r)u'(C_{ht})\right]$, where δ is the discount rate and r is the real interest rate on a single risk-free asset A owned by household h, and derive the key PIH equation under quadratic preferences, intertemporal separability, and perfect credit markets.

$\ln Y_{h,t-1} + D_{ht} + \varepsilon_{ht}$ where D_{ht} is a temporary government transfer and ε_{ht} is some non-forecastable *i.i.d.* component of labor income.

Using equation (2.1), we consider several possibilities for how expenditure patterns evolve over the course of a temporary cash transfer program. We consider two general regimes. In the first, households are informed at time $t-1$ about transfer disbursements over subsequent periods. In the second, households are not informed about these transfers. Using these two regimes, one can fully characterize the consumption behavior throughout the course of a given program subject to whatever data constraints the researcher faces.

Anticipated Beneficiary Status. Suppose that at time $t-1$, households learn about their beneficiary status. This implies that expenditure growth among non-recipients can be written as

$$\Delta \ln C_{h't} = \left(\frac{r}{1+r}\right) \varepsilon_{h't}. \tag{2.2}$$

This equation also generalizes to growth through $t+1$ by which time the program will have ended. There are two cases to consider for transfer recipients. Suppose that all identified beneficiaries are told at time $t-1$ that they will receive two quarterly transfer disbursements by time t and four quarterly disbursements by time $t+1$. If recipients obtain two transfer disbursements D by period t, then

$$\Delta \ln C_{ht} = \left(\frac{r}{1+r}\right) \varepsilon_{ht}. \tag{2.3}$$

That is, on average, these households exhibit identical mean expenditure growth to non-recipients.[10] However, if recipients obtain only one transfer by time t,

$$\Delta \ln C_{ht} = \left(\frac{r}{1+r}\right) (\varepsilon_{ht} - D_{ht}), \tag{2.4}$$

where the $-D_{ht}$ term captures the "surprise" effect of not having received the sec-

[10] Again, the assumptions necessary for causal identification in a difference-in-difference setting ensure that recipient and non-recipient households draw from the same income distribution *ex ante* (i.e., $\mathbb{E}[\varepsilon_{ht}] = \mathbb{E}[\varepsilon_{h't}]$).

ond disbursement by the time anticipated *ex ante*. The delayed arrival leaves the household with insufficient liquidity or cash-on-hand in the week(s) just prior to being observed by researchers at time t. If all beneficiaries receive the four quarterly disbursements by time $t+1$, then expenditure growth should be identical on average across all groups in the population.

Unanticipated Beneficiary Status. We can also consider how the predictions change if eventual recipient households did not anticipate the program at time $t-1$. In this case, equation (2.2) would still hold as non-recipients do not experience any transfer income shocks. However,

$$\Delta \ln C_{ht} = \left(\frac{r}{1+r}\right)(\varepsilon_{ht} + 2D_{ht}) \tag{2.5}$$

for households that obtained two disbursements by time t, and

$$\Delta \ln C_{ht} = \left(\frac{r}{1+r}\right)(\varepsilon_{ht} + D_{ht}) \tag{2.6}$$

for households that obtained only one disbursement by time t. In this case, beneficiaries should exhibit higher expenditure growth than non-beneficiaries, but the delayed disbursement still leads to differential expenditure growth across beneficiary groups. This differential no longer exists when looking at growth from t to $t+1$.

Changes in Expectations between Periods. It is often the case that the researcher cannot observe households at the precise time $s \in (t-1, t)$ when information about future income flows arrives. That is, the transfers are unanticipated over the period $t-1$ to s but anticipated from s onward. In this situation, $\ln(C_{ht}/C_{h,t-1})$ could be decomposed as

$$\Delta \ln C_{ht} = \left(\frac{r}{1+r}\right)[\lambda(\ln Y_{hs} - \mathbb{E}_{t-1} \ln Y_{hs}) + \gamma(\ln Y_{ht} - \mathbb{E}_s \ln Y_{ht})] \tag{2.7}$$

where the λ and γ captures the relative importance of income shocks before and after the mid-period information shock, respectively, with $\gamma + \lambda = 1$. To fix ideas, suppose that (i) the baseline survey is conducted at time t, (ii) the transfer announcement is made just prior to s when the first quarterly disbursement arrives, (iii) transfers proceed over the remaining quarter, and (iv) the follow-up survey is conducted at $t+1$. For beneficiaries with a delayed second disbursement, expenditure growth would then be given by

$$\Delta \ln C_{ht} = \left(\frac{r}{1+r}\right)[(\lambda-\gamma)D_{ht} + \lambda\varepsilon_{hs} + \gamma\varepsilon_{ht}] \tag{2.8}$$

and for those without any delay

$$\Delta \ln C_{ht} = \left(\frac{r}{1+r}\right)(\lambda\varepsilon_{hs} + \gamma\varepsilon_{ht} + \lambda D_{ht}) \tag{2.9}$$

Thus, even if the expenditure response to positive unanticipated transitory shocks is small (low λ), the negative shock associated with delayed disbursement may still lead to a large drop in expenditures if γ is high. Allowing for these asymmetries is especially important in the presence of credit constraints.

Credit Constraints. The framework above delivers testable implications for evaluating the expenditure impacts of cash transfer programs with well-publicized disbursement schedules and unexpected administrative delays—features not uncommon to large-scale transfer programs in developing countries. However, it is important to note that when r is small, all of the above equations imply a limited expenditure response to transitory income shocks so long as there are perfect credit markets. Before turning to our empirical tests, we briefly discuss the implications of relaxing this assumption given its limited applicability in a developing country context.

If the disbursements are anticipated but households cannot borrow upon learning of the future transfer income stream, then consumption will grow substantially at the time cash benefits are received. However, because households can still save, expenditures will respond much less to an anticipated decline in transitory transfer income. The opposite is true if the (delayed) disbursements are unanticipated. In this case, a negative shock to transfer income will lead to a sizable drop in expenditures in the absence of borrowing options whereas a positive shock will lead to a smaller change in consumption as households may use the transfers to engage in precautionary savings.

2.3 Empirical Tests

We test the model of consumption smoothing using a large-scale unconditional cash transfer (UCT) program in Indonesia. This temporary program affords us several sources of variation in transfer income not typically available in other cash transfer programs in developing countries. In this section, we provide relevant background on the program, describe our primary household panel dataset, and detail the empirical strategies for exploiting these rich sources of treatment variation.

2.3.1 Background on the UCT Program

In the midst of escalating global oil and gas prices in 2005, the Government of Indonesia (GoI) slashed fuel subsidies, raising regulated prices by a weighted average of 29 percent in February and then again by 114 percent in September. Although the subsidies were regressive, fuel products constitute a small share of overall household expenditures among both rich and poor.[11] However, the regulated increase in fuel prices led to generalized inflation as the the CPI grew by 17.9 percent between February 2005 and February 2006.

With the fiscal savings generated by the subsidy cutbacks, the government launched a temporary UCT program beginning in late 2005. The stated goal of the program—first announced publicly in August 2005—was to provide four quarterly disbursements of 300,000 Rupiah (Rp) (around 30 USD) to the poorest 30 percent of households beginning on October 1st. The transfers eventually reached every village in the country and were provided to households via local post offices, which were tasked with disbursing payments on scheduled dates for each village. Typically, such offices are located in subdistrict capitals and serve multiple villages.[12]

[11]Based on nationally representative household survey (*Susenas*) data from July 2004, households in the poorest decile of households allocate 3.7 percent of total monthly expenditures to kerosene on average while households in the richest decile spend only 1.9 percent. The figures for gasoline and diesel are similarly low but relatively more important for rich than poor households.

[12]In our data, described below, households report an average cost of 75 cents to reach the post

The targeting of beneficiaries proceeded in three stages. First, local government officials devised a large list of potential recipient households in August 2005 using a combination of own-discretion and community-based records from prior government programs. Second, using a minimalist survey instrument, the regional statistical bureaus enumerated households on the initial list as well as others from additional government sources.[13] Lastly, the Central Statistics Bureau (BPS) used the survey data to implement a proxy means test to generate the final list of eligible households by the end of September. Although the raw survey data and PMT scores are not available, the baseline *Susenas* data, which we describe next, include sound proxies for most questions.

2.3.2 Data

We use three waves of nationally representative panel data from the National Socioeconomic Survey (known as *Susenas*) collected in February/March 2005, 2006 and 2007. Matching households across the 2005 and 2006 rounds, we obtain a balanced panel of 9,048 households. We also observe a subset of households ($N = 7,016$) again in early 2007.[14] As discussed in Section 2.4.1, both the short- and medium-run panels have a large enough sample size to detect even small expenditure impacts.

Due to spatial variation in the timing of UCT disbursements, we observe three levels of treatment denoted by the number of disbursements $D \in \{0, 1, 2\}$ received by the time of *Susenas* enumeration in February and March 2006. There are 2,444 households in the treatment group ($D > 0$), but 639 of these households

office. However, reported costs range from zero to 10 USD. In some villages, local officials would arrange group transport and accompany households to the post office. In others, local officials would deliver disbursements directly to beneficiaries in the village.

[13] The survey questions concerned: (1) floor type, (2) wall and roof type, (3) toilet facility, (4) electrical source, (5) cooking fuel source, (6) drinking water source, (7) frequency of meat consumption, (8) frequency of meal consumption, (9) frequency of purchase of new clothes, (10) access to public health facilities, (11) primary source of income, (12) educational attainment of household heads, (13) amount of savings and type of assets, and (14) floor width.

[14] The baseline survey contains 10,574 households, while the follow-up in 2006 contains 9,892 households. The 2007 survey meanwhile contains more than 55,000 households, a subset of which were interviewed in preceding years. See Appendix 2.8.5 for details on panel construction, attrition, and sampling design.

had received a single disbursement at the time of enumeration while the remaining 1,805 households had received two disbursements.

Although progressively targeted, the UCT benefits did not reach many (near-)poor households. In Table 2.1, we find that recipient households are indeed poorer on average than non-recipients in early 2005 prior to the UCT program. However, there is strong evidence of potential (i) *leakage* of benefits as 37 percent of UCT recipients are in the top three national per-capita expenditure quintiles, and (ii) *undercoverage* as half of the lowest quintile did not receive any benefits. Figure 2.1 bears out these targeting results. Ultimately, only 50 (39) percent of poor (near-poor) households received any transfers despite being the nominal target population. We show next why these targeting errors make it possible to construct a credible group of counterfactual non-recipients in the absence of the PMT scores that were supposed to have dictated targeting.[15]

2.3.3 Identification

We are interested in the average treatment effect on the treated (ATT) of receiving d relative to s disbursements. Denoting this estimator by $\tau_{ds} \equiv \mathbb{E}[Y_h(d) - Y_h(s) | D_h = d]$ for some outcome Y_h for household h, we aim to identify the parameter vector $\boldsymbol{\tau} \equiv (\tau_{10}, \tau_{20}, \tau_{21})$ using the following difference-in-difference specification for the change in log consumption,

$$\Delta \ln C_{ht} = \kappa + \tau_{10}\mathbf{1}\{D_h > 0\} + \tau_{21}\mathbf{1}\{D_h = 2\} + \Delta\eta_{ht}, \tag{2.10}$$

where $\tau_{20} \equiv \tau_{21} + \tau_{10}$, η_{ht} is the idiosyncratic error term, and the constant κ is average growth among non-recipients. The $\boldsymbol{\tau}$ vector maps directly back to the predictions in Section 2.2. Our goal here is to ensure that the comparison of growth

[15] While it is not possible to obtain the actual administrative PMT scores that would enable a regression discontinuity design, we can use the available questions in *Susenas* coupled with the district-specific coefficients for each qualifying criteria to construct a strong approximation to each household's actual PMT score. However, as we show in Appendix 2.8.1, these reconstructed scores (i) fail to produce any fuzzy discontinuities around the stipulated threshold, and (ii) achieve less balance than our estimated propensity scores (see below) based on a richer set of household characteristics plausibly known to program enumerators and village officials at the time of beneficiary enrollment.

outcomes across groups is as close as possible to what one would observe if treatment status D had been assigned randomly.

Binary Treatment Effects. We pursue a reweighting approach in which the contribution of non-recipient households to the counterfactual is directly proportional to their estimated odds of treatment, $\widehat{\omega} = \widehat{P}/(1 - \widehat{P})$, where $\widehat{P}$ is the household's predicted probability of receiving *any* UCT benefits. Using a logit specification, we estimate this propensity score as an additive function of all observable underlying components of the PMT scores and additional household characteristics that would have been known to local targeting agents at the time of eligibility designation.[16] The full set of underlying parameter estimates are reported in Table 2.2, where we find that the likelihood of receiving any UCT benefits is: (i) higher for female-headed households, those benefiting from the long-running Rice for the Poor program, and households whose primary income source is in low-skilled occupations such as construction; (ii) decreasing in the size of land owned, housing floor area, and education level of the household head; and (iii) sensitive to housing status, the type of drinking water used, and toilet disposal location. The UCT program ultimately reached every village in the country, and overall, our propensity score model explains around one quarter of the variation in treatment status.

Figure 2.2 demonstrates the considerable overlap in propensity scores for treatment ($D > 0$) and control ($D = 0$) households. We can use the $\widehat{\omega}$ terms as inverse probability weights (IPW) in order to rebalance recipient and non-recipient households along observable dimensions. Under the assumption that there are no time-varying unobservable determinants of consumption growth correlated with UCT receipt, we can then assign a causal interpretation to the conventional binary treatment effect (see Abadie, 2005). In our case, this conventional ATT estimate is simply $\tau^{\text{binary}} \equiv \tau_{10} + \pi\tau_{21}$, where π is the share of recipients with $D = 2$.

Multivalued Treatment Effects. However, in order to identify the causal multivalued treatment effects parameters, $\boldsymbol{\tau}$, in equation (2.10), we must verify the

[16] The additive specification is in keeping with the weighting procedure used to estimate the original PMT scores.

exogeneity of the staggered rollout of the second quarterly disbursement. Table 2.3 shows that the probability of receiving disbursement two conditional on receiving disbursement one, $\mathbb{P}(D = 2 \mid D > 0)$, is explained almost entirely by geographic fixed effects. In columns 1-6, we find that household-level characteristics explain considerable variation in the probability of receiving any disbursements, $P(D > 0)$, even with more than 600 subdistrict fixed effects. However, turning to columns 7-12, household-level characteristics explain very little variation in $\mathbb{P}(D = 2 \mid D > 0)$ after controlling for district or subdistrict fixed effects. The R-squared and F tests in Table 2.3 suggest that the staggering occurs largely across (sub)districts and is plausibly exogenous with respect to baseline household characteristics. This is reassuring given that the quarterly disbursements were delivered via post offices typically located in subdistrict capitals.

We go a step further in Table 2.4 to show that both fixed and time-varying geographic characteristics cannot explain the spatial variation in staggering. Although district population size and the presence of banks explains some of the staggering process, relatively poorer or more remote regions do not receive the second disbursement any later than relatively wealthier, more central regions. Interestingly, distance to post offices has no predictive power. Nor does political affiliation of village officials with the central government in Jakarta. We also find that the actual date of survey enumeration in early 2006—days since the first household was enumerated on February 2—cannot explain the staggering process either. In other words, households waiting for their second disbursement at the time of enumeration were not simply residing in regions enumerated at earlier dates.[17] However, villages that experienced any mudslides or earthquakes in the previous three years are slightly less likely to have received the second disbursement by enumeration in early 2006. Nevertheless, the significant results in Table 2.4 should be interpreted cautiously as we expect that at least a few coefficients (5-10 percent) would be precisely estimated even if there were no relationship with staggering. Overall, Table 2.4 suggests that the staggering occurred for largely arbitrary administrative reasons and hence can be used to identify multivalued

[17]In fact, UCT recipients enumerated earlier were more likely to have received the second disbursement (see Figure 2.3).

treatment effects.[18]

Reweighting, Quasi-Random Staggering, and Balance. In Figure 2.4, we compare the distribution of baseline log household expenditures per capita across treatment levels. Given the exogeneity of the staggering process, it is not surprising that the distributions for treatment groups $D = 1$ and $D = 2$ are nearly identical and, in fact, statistically indistinguishable. Consistent with the summary statistics in Table 2.1, the control group is substantially richer at baseline than both treatment groups. However, once we reweight control households using $\widehat{\omega}$, the control group distribution shifts dramatically leftward and now effectively overlaps with the treatment groups' distributions. The slight imbalance in the extreme upper tail leads to a small albeit statistically significant mean difference across the treatment and control groups (at the 10% level). However, this poses a potential source of bias only insomuch as it cannot be explained by time-invariant determinants of consumption. Moreover, as demonstrated below, key results are robust to trimming this upper tail of $\ln C_{h,t-1}$. Other baseline covariates are effectively balanced after reweighting by $\widehat{\omega}$.[19]

Parallel Trends. We can also provide a partial test of the parallel trends assumptions underlying our identification strategy. Although we are unable to examine pre-program expenditure trends at the household level, we are able to do so at the district level using a representative estimate of average household expenditures from *Susenas* 2004 enumerated seven months prior to our baseline data in early 2005. We find that households with one disbursement by early 2006 are no more likely to reside in low-growth districts before the UCT program than are households with two or no disbursements.[20] This indirect evidence supports the parallel

---

[18] We do not consider other approaches to identifying multivalued treatment effects using the generalized propensity score (see Cattaneo, 2010; Imbens, 2000) since our multivalued treatment is plausibly exogenous. Augmenting the binary propensity score equation (see Appendix 2.8.1) with the covariates in Table 2.4 does not affect any of the key results below. To retain the full sample size, we omit these covariates in the main specifications.

[19] Empirically, less than 5 percent of the covariates in Table 2.2 exhibit statistically significant mean differences across recipients and non-recipients after re-weighting by $\widehat{\omega}$. These results are available upon request.

[20] In particular, we estimate the following equation by multinomial logit reweighting control

trends assumptions underlying the difference-in-difference identification and randomness of the staggering process.

Variation in Transfers per Capita. Because all households received the same transfer amount per disbursement, we observe considerable variation in transfers *per capita*, allowing us to identify an auxiliary measure of the marginal propensity to consume (MPC) out of transfer income. Figure 2.5 plots the distribution of transfers per capita in early 2006 for all recipients. In order to exploit the this intensive margin of treatment variation conditional on the number of disbursements d, we must first address two potential concerns. First, if the UCT program caused changes in household size, then any observed effect of $transfers/capita$ on expenditures may reflect this intermediate relationship. Second, local officials in some regions extracted a portion of the officially mandated 300,000 Rupiah disbursement per beneficiary. [21] If the incidence of informal taxes varied systematically depending on household size or other characteristics, then the estimated elasticity of expenditure growth with respect to transfers per capita might be biased. In Appendix 2.8.2, we rule out both of these concerns, demonstrating that (i) program receipt had no effect on household size, and (ii) the probability of recipient household h being taxed is orthogonal to a large array of household characteristics.

households by $\widehat{\omega}$:

$$P(D_{hjt} = d) = \alpha_d \ + \ \beta_d \Delta \ln \overline{C}_{j,t-1} \ + \ \upsilon_{hjt},$$

where D_{hjt} is the number of UCT disbursements d received by household h residing in district j in early 2006, and $\Delta \ln \overline{C}_{j,t-1}$ is the log growth in average household expenditures per capita in district j between July 2004 and early 2005. Computing marginal effects (MFE), we find that $\widehat{\beta_1^{MFE}} = 0.015$ ($se_1 = 0.038$) and $\widehat{\beta_2^{MFE}} = -0.027$ ($se_2 = 0.051$). This suggests that, after reweighting to achieve balance, households in the two treatment groups are no more likely to reside in low (or high) growth districts before the UCT program than are households in the control group.

[21] Approximately 6.5 (8.5) percent of recipients were subject to these informal taxes at the time of obtaining their first (second) UCT disbursement. According to these recipients, the proceeds covered local administration, security at disbursement location, but most were intended for redistribution to non-recipients deemed deserving by local officials.

2.4 Cash Transfers and Expenditure Growth

We turn now to the main empirical results aimed at testing the hypothesized consumption smoothing behavior. In addition to pure OLS, we consider four alternative reweighting estimators of equation (2.10). All are predicated on the *inverse probability weighting* (IPW) approach of reweighting control households by their estimated odds of treatment, $\widehat{\omega}$. The *double robust* estimator augments the IPW specification with the propensity score ($\widehat{P}$) or the covariates ($\mathbf{X}_h$) used to predict those scores. The *heterogeneous control function* estimator introduces a fifth-order polynomial in the propensity scores and allows it to vary across recipients and non-recipients.[22] We trim all households with estimated propensity scores $\widehat{P} > 0.91$, which is the optimal bound using the Crump et al. (2009) procedure. In all specifications, we control for province fixed effects and cluster standard errors at the village level, using a block bootstrap whenever the generated propensity score is used.[23]

2.4.1 Main Results: Timing Matters

We begin by presenting estimates of equation (2.10) that are consistent with the key predictions in Section 2.2. The top panel in Table 2.5 reports baseline results for the growth in log total household expenditures per capita between 2005 and 2006. We find a consistent pattern of differential treatment effects across all reweighting specifications: Recipients still awaiting their second disbursement at the time of enumeration in early 2006 have significantly lower expenditure growth relative to non-recipients *and* recipient households with both disbursements. In the

[22]This approach is akin to a Oaxaca-Blinder decomposition and preserves the benefits of double robustness (see Kline, 2011). Results are unchanged when using other polynomial orders. An excellent review of reweighting estimators can be found in Busso et al. (2009) and Imbens and Wooldridge (2009). We prefer these estimators to the more common matching procedures used in the evaluation literature because (i) reweighting estimators often have better finite sample properties than standard matching estimators in situations with considerable overlap in propensity scores as we have here, and (ii) these estimators make it computationally easy to recover *multivalued* treatment effects.

[23]More specifically, we (i) draw all households from a random sample of the 629 villages, (ii) recover $\widehat{P}$ and $\widehat{\omega}$ using a logit specification, and (iii) estimate $\boldsymbol{\tau}$ in equation (2.10) using the given reweighting estimator and clustering standard errors by village. With over 600 villages, there is little concern about small cluster bias (see Cameron et al., 2008b).

most flexible control function specification in column 5, these single disbursement recipients have expenditure growth rates that are 0.076 log points lower than non-recipients. Moreover, expenditure growth among recipients of two disbursements is statistically indistinguishable from the average of 0.11 among non-recipients. The null results are fairly precisely estimated as the standard errors on τ_{20} are roughly 5 percent of one standard deviation of $\Delta \ln C_{ht}$. The results are largely insensitive to the estimator used with the exception that the OLS estimates of τ_{10} and τ_{21} are slightly lower.[24] However, had we pooled the two recipient groups and merely estimated a conventional binary treatment effect (τ^{binary}, see Section 2.3.3), we would have understated by nearly 30 percent the expenditure gains to receiving the full two transfers as expected by early 2006.

Retaining the same specifications and moving ahead to early 2007, the bottom panel of Table 2.5 shows that the differential treatment effects dissipate over the two-year time horizon by which time all recipients had obtained the full four quarterly disbursements. Both groups of treatment households are statistically indistinguishable from control households.[25] Moreover, these null results are again fairly precisely estimated.

To be sure, the sample sizes are sufficiently large to identify small impacts as the data allow for minimum detectable effects (MDE) of around 0.05 standard deviations. This calculation is based on (i) a sample of 9,048 households, (ii) baseline variance of log expenditures of 0.56, and (iii) one-third of the population receiving the treatment. The implied MDE compares favorably to the standard deviation of observed log expenditure per capita growth (roughly 0.4). Similarly small MDEs are found for the smaller sample over 2005–2007. Given that we use monthly expenditures and each quarterly disbursement comprises nearly 45 percent of average monthly expenditures in recipient households at baseline, we could observe treatment effects as large as one standard deviation of growth if

[24] Yet, the OLS estimates are statistically indistinguishable from the reweighting estimates. One explanation could be that selection bias is limited after taking first-differences and hence may be largely confined to the cross-section. However, in each of the columns 3-5, the selection terms are (jointly) statistically significantly different from zero.

[25] These estimates are not an artifact of the attrition of households between 2006 and 2007 survey rounds (see Section 2.3.2). Key results remain largely unaffected when reweighting the sample to account for the probability of attrition.

households consumed all of the transfers within the enumeration period.

If disbursement delays cause sizable reductions in expenditures, then the impact of the shock should likely be increasing in the length of the delay as cash-in-hand constraints become more binding. Table 2.6 provides indirect evidence of this mechanism by showing that the drop in expenditures among recipients still waiting for the second disbursement is larger for households enumerated in March than for those in February 2006.[26] In fact, average expenditure growth among those enumerated in February 2006 is indistinguishable from non-recipients and those that already received a second disbursement. The similarity of τ_{10}^{March} with τ_{10} in Table 2.5 suggests that the average negative expenditure impact of realized transfers falling short of expectations is likely driven by those households experiencing the longest delay. Before further relating the estimates of $\boldsymbol{\tau}$ in these tables to the predictions in Section 2.2, we briefly discuss a few results that shed additional light on the timing of consumption.

Decomposing Expenditure Growth. Table 2.7 shows that the differential treatment effects are driven largely by changes in expenditures on food rather than non-food items. Using the flexible control function estimator, we cannot reject the null hypothesis that all three groups $D \in \{0, 1, 2\}$ have identical non-food expenditure growth over both the short- and medium-term periods. Over the short-term period, the key parameters τ_{21} and τ_{10} are nearly halved when restricting attention to non-food expenditures. However, the opposite is true for food expenditures, where we find that these parameters are amplified and statistically significant at the one percent level.[27] Although the differences in coefficients between columns 1 and 2 are not statistically significant, these results suggest that the expenditure growth differential between recipient groups can be attributed primarily to differences in food expenditures over the week prior to survey enumeration in early

[26] Around one quarter of households were enumerated in February with a proportional share of recipients still awaiting the second quarterly disbursement.

[27] Angelucci and Attanasio (forthcoming) show that the *Oportunidades* cash transfers led to larger increases in food than non-food expenditures among poor households in urban Mexico. In Appendix 2.8.2, we find that the effects of the UCT do not differ in urban relative to rural areas of Indonesia where households grow more of the food they consume.

2006.

In Table 2.8, we further disaggregate food and non-food expenditure items and find the same general patterns.[28] For most expenditure subcategories, (i) recipients still awaiting their second disbursement have lower expenditure growth than recipients of two disbursements and non-recipients, and (ii) the second disbursement closes the gap between recipient and non-recipient expenditure growth. However, we find no effect of the transfers on purchases of grain including rice and tubers. Given the storability of grains, it is likely that such purchases are less sensitive to the cash-in-hand mechanism underlying the treatment effects for other, more perishable foods. In other words, grains are the least likely to have a one week shelf-life that would have led to their entry in the expenditure survey module. For reasons we discuss next, these expenditure decompositions provide further evidence that timing matters.

Mapping $\boldsymbol{\tau}$ to Theory. At first glance, the pattern of coefficients in Tables 2.5-2.8 seem to be at odds with a large literature documenting sizable expenditure gains to transfer programs in low-income settings. However, as we discuss now, these baseline findings are largely consistent with the theory of consumption smoothing developed in Section 2.2.

If we define $t-1$ as the period immediately after the announcement of the program benefits and implementation schedule, then the treatment effects reported in Table 2.5 are consistent with equations (2.2)-(2.4). That is, (i) $\tau_{20} \approx 0$ captures the difference between equations (2.3) and (2.2), (ii) $\tau_{10} \approx -0.075$ captures the difference between equations (2.4) and (2.2), and (iii) $\tau_{21} \approx 0.075$ captures the difference between equations (2.3) and (2.4). These model predictions are also consistent with the largest expenditure differences being observed for (i) households experiencing the longest delays, and (ii) (perishable) food rather than non-food items since the former is reported over the week immediately prior to enumeration

[28] In keeping with the specification for aggregate expenditure growth, we restrict the estimates for each commodity group to those households with non-zero expenditures in both periods. In unreported results, we find little evidence of any differential treatment effects along the extensive margin in terms of switching in or out of nonzero expenditures across the categories of goods in Table 2.8 (results available upon request).

whereas the latter is reported over the month prior to enumeration.

Moreover, this framework can also explain why the differential treatment effects in Table 2.5 dissipate by 2007 (i.e., $\tau_{20} = \tau_{10} = \tau_{21} \approx 0$). First, the surprise shock in equation (2.4) no longer holds as all recipients received the four quarterly transfers as expected by the end of 2006 ahead of enumeration in early 2007. Second, even if credit constraints were binding, households can engage in *ex ante* savings (out of the transfers) ahead of the anticipated end of the program in late 2006.

Although the key estimates seem to be consistent with the theory of consumption smoothing under anticipated beneficiary status, we only observe consumption in early 2005 several months prior to the public announcement of the UCT program. It is therefore important to consider how our interpretation changes under the strict assumption that households did not anticipate their beneficiary status at the time of baseline enumeration. In this case, $\tau_{21} > 0$ is still consistent with the model in equations (2.5) and (2.6), but $\tau_{20} \approx 0$ and $\tau_{10} < 0$ are not. The unanticipated positive income shock should have led to higher expenditure growth among recipient households (i.e., $\tau_{20} > 0$ and $\tau_{10} > 0$). However, if we incorporate the sharp break in expectations of future income with announcement of the UCT program, equation (2.7) suggests a potential reconciliation between the two timing regimes. In particular, given that we are measuring short-term expenditures over the month prior to enumeration in early 2006/7, it is possible that the negative shock associated with the delayed second quarterly disbursement has a much larger effect on the observed change in consumption than the positive shock associated with the arrival of the first quarterly disbursement around October 2005. This could explain why we find that $\tau_{10} < 0$ (and $\tau_{20} \approx 0$) and is a particularly relevant interpretation if liquidity constraints are binding—an issue addressed formally below.

Until now, we have ignored the possibility that poor non-recipients in the control group may have expected to benefit from the UCT program, which seems likely given the pervasive targeting errors. However, even if this were true for all potentially eligible beneficiaries in the control group, their expenditure changes should work against what we observe empirically since these households experience a negative income shock that pushes down average growth in the control group.

It is important to note that our strict model-based interpretation of Tables 2.5-2.8 hinges on the real interest rate r being non-zero. Taking the model seriously, our estimates of τ_{21} from Table 2.5 imply $\widehat{r} \approx 0.075$, which seems somewhat high. Although nominal interest rates quoted were indeed quite high at this time, so was inflation on account of the fuel subsidy cutbacks. However, as discussed in Section 2.2, even if the true $r \approx 0$, households may exhibit large responses to transitory income shocks if they are liquidity-constrained. We turn now to a formal investigation of the potential ways in which these constraints might affect our results and theoretical interpretation.

2.4.2 Financial Institutions Moderate the Expenditure Response to Transfers

There are several reasons why credit constraints can lead to departures from strict PIH behavior. In this section, we show how the expenditure response to the UCT program may depend on credit access. For every village in the *Susenas* data, we identify a "bank nearby" if there are any banking institutions operating within the village's subdistrict in early 2005.[29] In our context, proximity to banks serves as a reasonable proxy for the liquidity of local credit markets which might facilitate consumption. We investigate this possible interaction with the program by augmenting equation (2.10) with an indicator for bank presence and its interaction with treatment indicators.

Overall, the results in Table 2.9 imply that the presence of financial institutions moderates the expenditure response to cash transfers. First, the negative short-term expenditure shock from experiencing a late second disbursement is almost entirely offset by residing near banks. The null coefficient τ_{10}^{banked} suggests

[29] The measure is based on the presence of any formal banking institutions reported in the Village Potential (or *Podes*) data from 2005 (see Appendix 2.8.5). Around 75 (80) percent of treatment (control) households reside close to these banking institutions. There is a measure of household-level credit access in *Susenas*, but it is limited to a very specific type of credit, use of which is too limited for econometric purposes (less than 1.5 percent of households report any use in 2005).

that expenditure growth for single disbursement recipients near banks is statistically indistinguishable from non-recipients.[30] At the same time, the difference between recipient groups also dissipates among those UCT beneficiaries located near banks. Second, proximity to banks seems to have no differential effect on UCT recipients that obtained the second disbursement early in 2006 (compare τ_{20}^{banked} and $\tau_{20}^{unbanked}$). This is intuitive since, on average, these households did not experience any surprise shocks in response to which they would have needed to draw on local financial markets. Last, by early 2007, we observe that UCT recipients residing near banks have slightly higher expenditure growth than recipients residing farther away from banks ($\tau_{d0}^{banked} > \tau_{d0}^{unbanked}$ for $d = 1, 2$). There are surely differences in local economic vitality and expenditure growth trends that are correlated with bank presence. Hence, this final result in Table 2.9 could be explained by banked and unbanked recipient households returning to their differential pre-program expenditure trends several months after receiving the final UCT disbursement.[31]

2.4.3 Age Heterogeneity in the Expenditure Response to Transfers

There are two interesting implications of the life cycle version of the PIH model that are, in principle, also (indirectly) testable in our setting. First, in a setup with natural borrowing (i.e., budget) constraints, the consumption response to transitory income shocks should be (slightly) increasing in age as individuals near retirement at the end of the life cycle. Second, in a setup with zero borrowing constraints and a non-negative asset condition, younger households with limited savings should exhibit larger responses to transitory income shocks. Blundell et

[30] These null results are again fairly precisely estimated zeros.

[31] We also explore whether household members working abroad may be another vehicle for smoothing consumption *ex post*. Yang and Choi (2007), for example, finds that remittances can act as insurance against negative rainfall shocks in the rural Philippines. In our setting, households still awaiting their second quarterly disbursement may have been able to draw on remittances from family members abroad to smooth the negative income shock. In results available upon request, we take a similar approach as in Table 2.9 (replacing bank nearby with an indicator for migrant abroad) and find some evidence for this mechanism, but the results are not robust.

al. (2008) and Kaplan and Violante (2010) provide new evidence leading to this U shaped age profile of the marginal propensity to consume (MPC) out of income shocks. To capture these potential nonlinearities, we augment equation (2.10) with a quadratic in the age of the household head and the interactions of those terms with the two treatment indicators.

The key results of this exercise are summarized succinctly in Figure 2.6, which plots the marginal effects, $\boldsymbol{\tau}$, at every age in the data. Although the estimates are imprecise (see Appendix 2.8.2), we find an age profile of treatment effects that is consistent with the predicted convex relationship. In particular, the parameters τ_{10} and τ_{20} are largest for the youngest and oldest Indonesian households.

2.4.4 Robustness, Spillovers, and Rainfall

In this section we show that the main empirical results (i) hold up to a number of robustness checks, (ii) cannot be explained by systematic spillovers to the control group via local redistribution, and (iii) are consistent with household responses to other transitory income shocks due to rainfall.

Robustness Checks. In Table 2.10, we illustrate a battery of robustness checks using the control function estimates for the period from 2005–2006 as a baseline.[32]

Timing of the Midline Survey. Our identification strategy relies on differences in the disbursement schedule across households enumerated at roughly the same point in time in early 2006. Using a coarse indicator for later enumeration, we saw in Table 2.6 that the negative shock experienced by recipients with delayed disbursements is larger for those enumerated in March than in February 2006. We

[32]In unreported results, we also consider alternative estimators for the binary treatment effect of receiving any UCT benefits including nearest-neighbor matching (Abadie and Imbens, 2005), local linear matching (Heckman et al., 1998), and inverse probability tilting (IPT) (Graham et al., 2012). In all cases, the main qualitative and quantitative findings remain unchanged from those binary treatment effect estimates recoverable from Table 2.5. Moreover, as we show in Appendix 2.8.2, there does not appear to be heterogeneity in the multivalued treatment effects across the distribution of expenditure growth (i.e., the effects at the mean are statistically indistinguishable from the effects at other quantiles).

further ensure that differential enumeration dates are not driving our results by including exhaustive dummies for the 65 distinct days of enumeration across the country. Doing so in row 2 of Table 2.10 leaves the results unchanged from the baseline estimates in row 1.

Alternative Geographic Fixed Effects. Row 3 shows that results are robust to including district fixed effects as well as to clustering standard errors at any administrative division above the village. However, as expected based on the results in Table 2.3, including subdistrict (village) fixed effects in row 4 (5) removes nearly all of the exogenous variation in the staggering of the second quarterly transfer and pushes the estimates closer to a simple binary treatment effects specification.

Trimming Extreme Expenditures. In rows 6 and 7, the key qualitative results do not change if we trim (i) the top and bottom percentile of $\Delta \ln C_{ht}$, or (ii) the top and bottom percentile of $\ln C_{h,t-1}$.

Regional Differences in Inflation. By including province fixed effects, we remove trend differences across regions in terms of inflation and hence of the passthrough from fuel price increases to other consumer goods. We take two additional steps in rows 8 and 9 to ensure that local price differences are not driving our results. First, we deflate nominal expenditures using the nearest official regional CPI measures (see Appendix 2.8.5). Second, we deflate using the price of the goods basket used to construct the district-specific poverty lines. Neither approach materially affects our estimates of $\boldsymbol{\tau}$.

Alternative "per capita" Formulations. Some authors argue that when looking at household expenditure outcomes, one should account for the fact that children require less consumption than adults to attain equivalent levels of welfare (see Deaton, 1997b; Olken, 2006). To the extent that household composition differs across treatment and control groups (see Table 2.2), this could impact our results. Ultimately, though, this adjustment is irrelevant as the baseline estimates in row 1 are indistinguishable from those in rows 10 and 11 where we treat children aged 0-9 years old as 0.5 adult-equivalents for total and food expenditures, respectively.

Durable Goods Expenditures Beyond the Last Month. In the baseline regressions, we measure durable goods expenditures in the last month. In so doing, we may have missed important purchases using UCT funds prior to the survey enumeration period in early 2006. In other words, the UCT may have led to an increase in expenditures several months prior to enumeration and perhaps immediately after UCT receipt in October–December 2005. Hence our comparison of durable goods purchases in the early months of 2005 and 2006 might understate the large positive effects of the UCT had we compared those purchases going back over the full year prior to enumeration. This does not seem to be the case. Using a (pro-rated) measure of annual non-food expenditures in row 12 leaves our key parameter estimates unchanged. This provides some additional evidence that the expenditures response to positive and negative shocks is asymmetric. That is, the positive income shock associated with the first disbursement leads to a smaller expenditure impact than the negative shock associated with the delayed second disbursement.

Participation in Other Social Programs. Several other previously operative social programs continued alongside the UCT. Receipt of such programs might confound our estimates of $\boldsymbol{\tau}$ parameters if, for example, the UCT disbursement schedule was timed so as to reach those households lacking other programs first. In row 13, we control for participation in other programs—a rice subsidy scheme, scholarships for poor students, and subsidized health insurance for the poor—and the results remain similar to the baseline.

Idiosyncratic Health Shocks. In row 14, we show that the results are unchanged when conditioning on changes in the incidence of health shocks within the household between 2005 and 2006. This is reassuring given that health shocks are potentially important time-varying omitted variables correlated with both treatment assignment and expenditures.

Local Natural Disasters. In row 15, we control for the incidence of local natural disasters from 2003-5 and find no systematic departures from the baseline estimates of $\boldsymbol{\tau}$. This is reassuring given the slight correlation found in Table 2.4 between late disbursement and these other shocks to (income and hence) expenditures arising from natural disasters.

Other Covariates of Staggering in Table 2.4. We go a step further in row 16 and control for *all* covariates used to explain staggering in Table 2.4. Doing so leaves the main baseline findings largely unchanged. However τ_{21} falls slightly, suggesting that regional characteristics correlated with disbursement timing may explain some of the differential expenditure growth across recipients.

Program Enrollment and Systematic Underreporting of Expenditures. In the midst of public scrutiny over perceived program leakage and undercoverage, it is possible that UCT recipients and particularly those still awaiting their second disbursement systematically underreported their expenditures.[33] This would lead to non-classical measurement error and could bias the treatment effects downward if recipients perceived their ongoing participation as being contingent upon reported welfare levels. We (partially) test for this by controlling for whether the household was assigned to the initial list by the village head (potentially more prone to patronage) or by a regional government official outside the village (less prone to patronage). The drop in magnitude and significance of τ_{10} in row 17 of Table 2.10 relative to Table 2.5 provides some indirect evidence in support of this mechanism. Conditional on official program enumerator visits, there are no longer statistically significant differences between control households and recipients of a single transfer.

Ruling Out Spillover Effects. We can also provide indirect evidence that potential spillovers to control households do not explain the observed expenditure impacts observed in Table 2.5. We proceed in three steps. First, we identify villages in which UCT recipients report any informal taxes during disbursement round one or two.[34] Second, we assign all control households in these villages to treatment

[33] The first few months of the UCT program in 2005 generated considerable public controversy surrounding the allocation of benefits and widespread perception of mistargeting (see Cameron and Shah, 2012).

[34] Out of 538 villages with any UCT recipients, we find that 17 (16) report informal taxes during the first (second) disbursement period, and 23 report informal taxes in both periods. Among those taxed, the median amount also increased from 20,000 Rp to 50,000 Rp. The portion allocated to supposed local redistribution increased from 40 percent at the first disbursement to 62 percent at the second disbursement.

group $D = 1$ ($D = 2$) if any informal taxes are observed for disbursement one (two). Third, we re-estimate the propensity scores and finally the key parameters in Table 2.5. In doing so, we find that τ_{10} remains largely unchanged while τ_{21} and τ_{20} fall relative to the baseline estimates. This is the opposite of what we should observe if spillovers were contaminating the control group and hence raising their expenditures and systematically biasing our baseline treatment effects downward. Of course, these results (available in Appendix Table 2.18), should be interpreted with caution as (i) surely not all non-recipients benefited from informal redistribution of UCT taxes within given villages, and (ii) our test implicitly assigns equal transfers across households.

Other Transitory Income Shocks. Following others in the development literature beginning with Wolpin (1982), we examine the expenditure response to rainfall shocks as another test for excess sensitivity to transitory income shocks.[35] The key message from Table 2.12 is that these shocks are associated with higher growth in household expenditures per capita among rural households engaged in agriculture. In column 3, we find that a 10 percent deviation of rainfall from its long-run mean yields roughly a 1.8 percent increase in expenditures per capita. We find in column 4 a similarly positive elasticity for land-owning households, and in column 5 the elasticity is increasing in land-holding size—though only the interaction terms are significant in both cases.[36]

2.4.5 Intensive Margin Treatment Effects: An Auxiliary Estimate of the MPC

Before concluding, we provide an auxiliary estimate of the MPC out of transfer income by exploiting the fact that the transfer size per disbursement was fixed across households of varying size. This allows us to estimate the following

[35]The transitory rainfall shock in year t is defined as the log rainfall level in village v's district over the province-specific growing season minus the log mean rainfall level for that district over the forty years/seasons prior to t.

[36]Although rainfall shocks only affect the transitory income and hence expenditures of certain segments of the (rural) population, the UCT benefits do not have heterogeneous effects along these same dimensions (see Appendix 2.8.2).

equation:

$$\begin{aligned}\Delta \ln C_{ht} &= \alpha + \tau_{10}\mathbf{1}\{D_h > 0\} + \tau_{21}\mathbf{1}\{D_h = 2\} \\ &\quad + \psi \; transfers/capita_{ht} + \sum_{j=2}^{13} \beta_j \mathbf{1}\{HHsize_{ht} = j\} + \Delta\eta_{ht}, \end{aligned} \tag{2.11}$$

where $transfers$ is the total amount of UCT funds (in 100,000s of Rupiah) received by enumeration in early 2006, and $capita$ and $HHsize$ are household size. After removing (i) the multivalued treatment effects through reweighting and the disbursement indicators, and (ii) the independent effects of household size through β_j terms, all that remains is information on the scale or intensity of UCT benefits. Under the assumptions verified in Section 2.3.3 as well as $\mathbb{E}[HHsize_{ht}\Delta\eta_{ht}] = 0$ (after reweighting), ψ identifies the marginal effect of an additional unit of transfers income per capita.

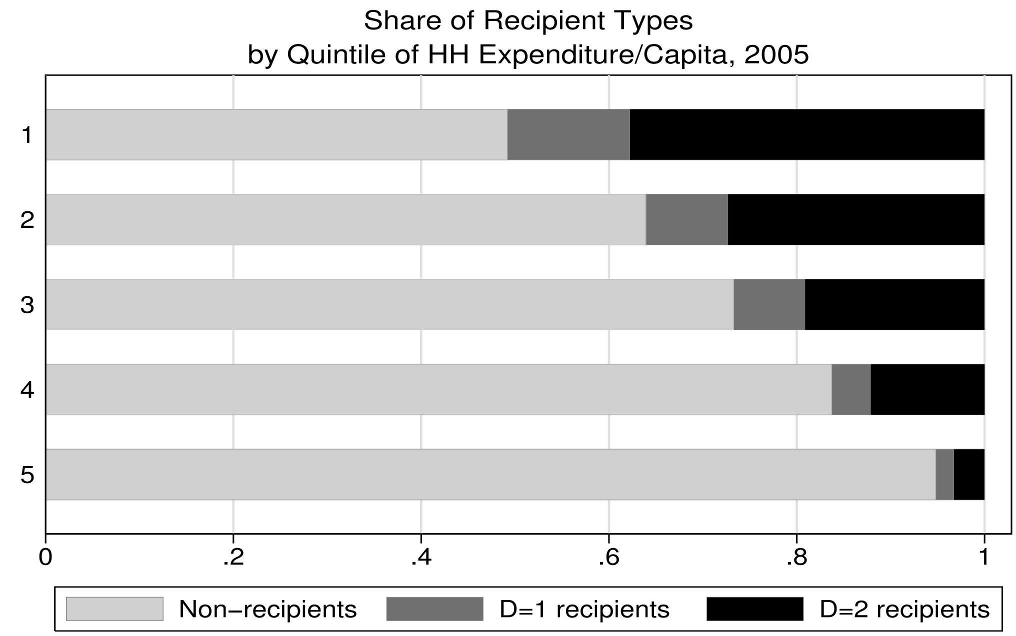
In Table 2.11, we find robust positive estimates of ψ. Columns 1-3 impose $\beta_j = 0$ for all j, and column 4 permits $\beta_j \neq 0 \; \forall j$ to allow for unconditional scale effects in the growth in household expenditures/capita. The point estimates of 0.04-0.065 for total expenditures per capita imply a MPC out of transfer income of around 0.08-0.11. This suggests that an increase in household transfers per capita by 10 USD per quarter—roughly 1/24th of household expenditures at baseline—yields roughly a 5 percent increase in monthly expenditures per capita. Going back to the theory, the estimated MPC implies a similar real interest rate r as obtained from the mulitvalued treatment effects. In the classical PIH setup, the size of income shocks does not matter. However, if adjustment costs are high and households are liquidity-constrained, this need not be true. Coupled with the evidence in Section 2.4.3, these results suggest that expenditures of credit-constrained households may be sensitive to size of shocks in a context with prohibitively high borrowing costs.

2.5 Conclusion

This paper has demonstrated the importance of incorporating timing and expectations into the evaluation of household expenditure responses to cash transfers in a developing country context. Using a simple model of consumption smoothing based on the permanent income hypothesis, we investigate the effects of a large scale, temporary UCT program on household expenditure growth in Indonesia. On average, beneficiary households that received the full two transfers as expected by early 2006 do not differ from comparable non-beneficiaries in terms of per capita expenditure growth. However, beneficiaries still unexpectedly awaiting their second transfer report significantly lower expenditure growth especially in areas with limited financial access. Using the third wave of panel data, we find that these growth differentials dissipate by early 2007, several months after the final transfer was received by all beneficiaries. Using the fact that the transfer amount per disbursement was fixed across households, we also identify an auxiliary marginal propensity to consume out of transfer income of around 0.1, which is consistent with our short-run differential treatment effects. Overall, the results suggest that households respond primarily to unexpected transitory changes in transfer income with the response to positive shocks being much smaller than the response to negative shocks.

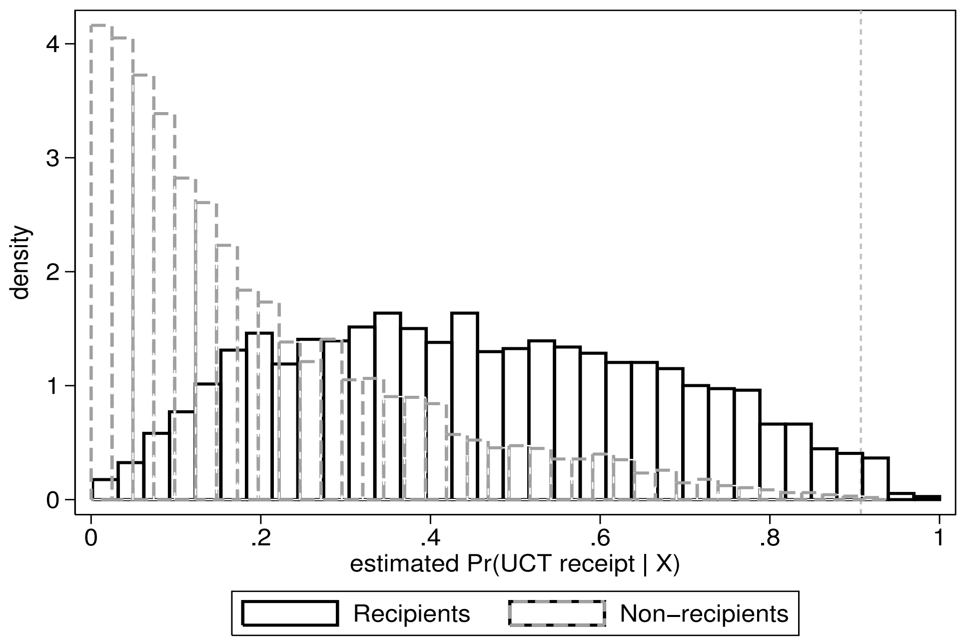
More broadly, our paper offers new insights into the expenditure impacts of large-scale fiscal interventions in a developing country context. Unlike numerous programs in Latin America and elsewhere, the UCT in Indonesia was not explicitly designed as a transformative poverty alleviation program. Rather, the government used the program as a means of transitioning away from regressive fuel subsidies. Similar subsidy reforms have either recently been implemented or are being considered in a number of developing countries (Coady et al., 2010). These programs have a number of important welfare implications and warrant further study. Our results from Indonesia suggest that the household response to cash transfers in such contexts may hinge on perceived program duration as well as the timing of the transfers with respect to subsidy cutbacks. In order to understand the full policy-relevant impact of these (and other) programs, evaluators must explicitly test for behavior change at various times before, during, and after the actual interventions.

2.6 Figures



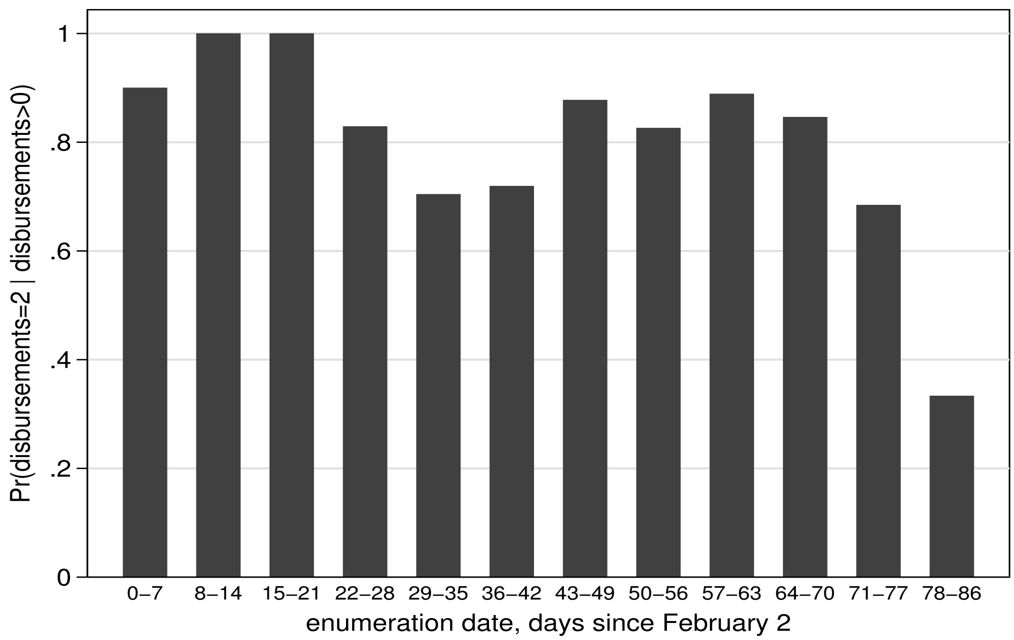
Notes: D denotes the number of UCT disbursements d received by enumeration in early 2006 as reported in a module attached to *Susenas* 2006. The quintile of household expenditures per capita is based on data reported in *Susenas* 2005.

Figure 2.1: Treatment Level by Baseline Expenditure Decile



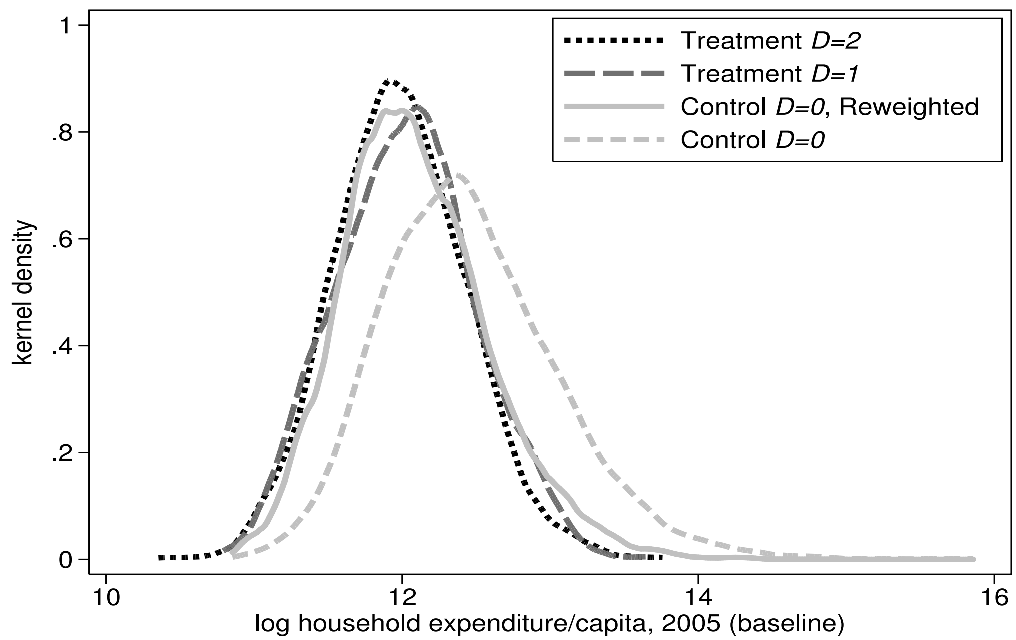
Notes: Propensity scores obtained from flexible logit regressions (see Table 2.2). Observations to the left of the dashed vertical line fall within the Crump et al. (2009) optimal overlap region.

Figure 2.2: Overlap in Estimated Propensity Scores ($\widehat{P}$)



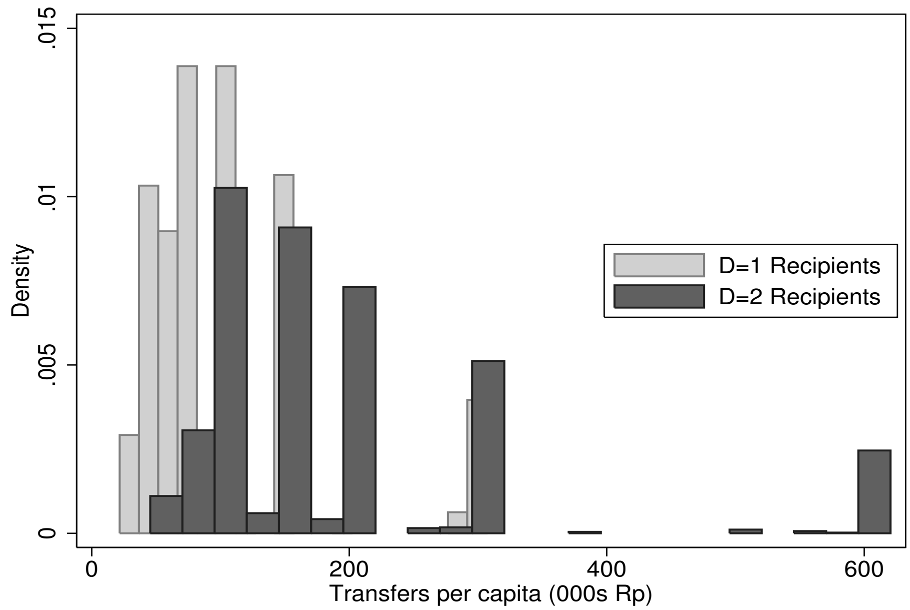
Notes: The bars represent the share of UCT beneficiary households with two disbursements at the time of being visited by *Susenas* enumerators in early 2006. The vast majority of enumeration dates fall between February 15 and March 25.

Figure 2.3: Staggered Disbursements and Survey Enumeration Date



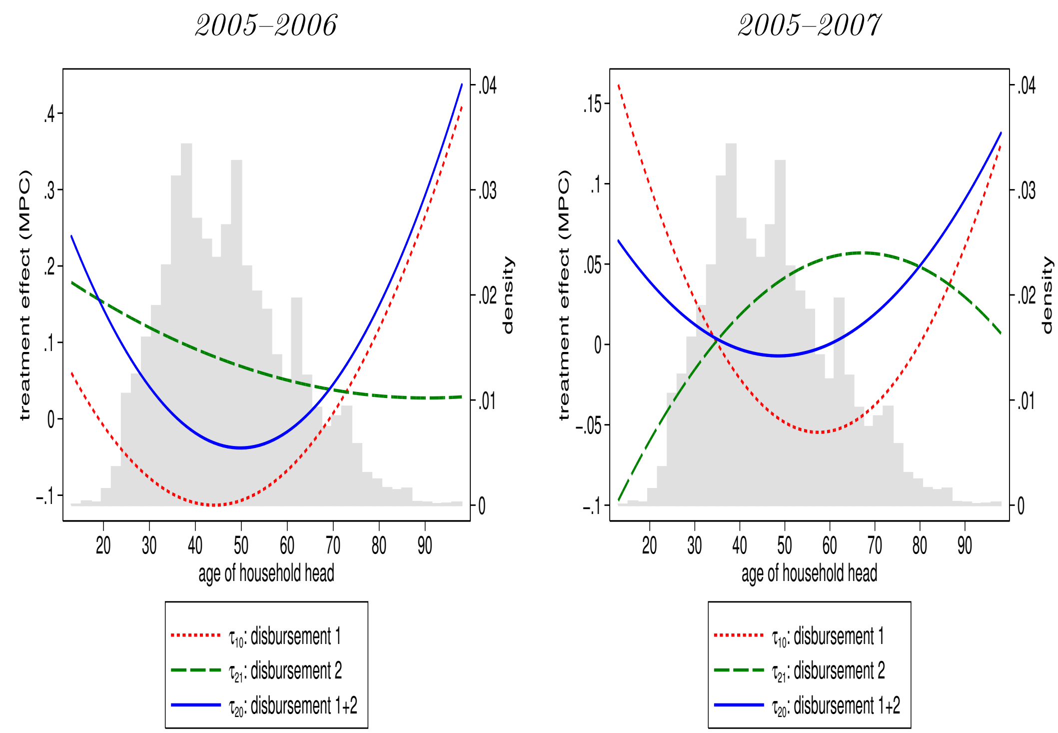
Notes: D denotes the number of UCT disbursements d received by early 2006. All distributions estimated using Epanechnikov kernel and a rule-of-thumb bandwidth. The "Control (Reweighted)" observations are adjusted using inverse probability weights (IPW) based on normalized estimated odds of treatment $\widehat{\omega} = \widehat{P}/(1-\widehat{P})$.

Figure 2.4: Baseline Expenditure Distributions by Treatment Status



Notes: D denotes the number of UCT disbursements d received. The transfer amount reported per disbursement is also reported *Susenas* 2006.

Figure 2.5: Distribution of Transfers per Capita through early 2006



Notes: Each curve represents the treatment effects coefficients estimated across household head ages. The marginal effects are recovered from a regression allowing the $\boldsymbol{\tau}$ parameters to vary with a quadratic in age.

Figure 2.6: Age-Specific Treatment Effects

2.7 Tables

Table 2.1: Expenditure statistics, 2005 and 2006

	2005					2006				
	Mean	**SD**	**Min**	**Median**	**Max**	**Mean**	**SD**	**Min**	**Median**	**Max**
	Non-recipients ($N = 6604$)									
Expenditure/capita (000s Rp)	315	292	52	243	7702	356	300	31	272	4891
Food expenditure/capita (000s Rp)	162	93	30	138	2790	182	104	20	155	1141
Non-food expenditure/capita (000s Rp)	153	234	8	94	7071	174	228	0	108	4236
Education expenditure/capita (000s Rp)	11	60	0	2	2269	8	41	0	0	1660
Health expenditure/capita (000s Rp)	11	67	0	2	2607	10	62	0	2	3137
Below poverty line	0.10	0.30	0	0	1	0.11	0.31	0	0	1
Quintile (nat'l) expenditure/capita	3.23	1.38	1	3	5	3.28	1.37	1	3	5
Quintile (intra-province) expenditure/capita	3.21	1.39	1	3	5	3.26	1.38	1	3	5
	$D = 1$ ***Recipients*** ($N = 639$)									
Expenditure/capita (000s Rp)	185	93	49	165	843	195	118	41	170	1817
Food expenditure/capita (000s Rp)	121	60	32	110	761	123	62	30	110	422
Non-food expenditure/capita (000s Rp)	65	49	9	52	423	72	80	9	56	1581
Education expenditure/capita (000s Rp)	2	6	0	0	220	2	4	0	0	48
Health expenditure/capita (000s Rp)	10	83	0	1	1832	4	11	0	1	150
Below poverty line	0.25	0.43	0	0	1	0.34	0.47	0	0	1
Quintile (nat'l) expenditure/capita	2.25	1.21	1	2	5	2.14	1.18	1	2	5
Quintile (intra-province) expenditure/capita	2.36	1.27	1	2	5	2.27	1.25	1	2	5
	$D = 2$ ***Recipients*** ($N = 1805$)									
Expenditure/capita (000s Rp)	178	90	31	159	945	192	92	37	172	908
Food expenditure/capita (000s Rp)	115	54	17	104	645	124	57	23	112	484
Non-food expenditure/capita (000s Rp)	63	50	9	50	576	68	51	0	55	682
Education expenditure/capita (000s Rp)	3	8	0	0.4	220	2	5	0	0	68
Health expenditure/capita (000s Rp)	5	11	0	2	178	5	22	0	1	751
Below poverty line	0.28	0.45	0	0	1	0.31	0.46	0	0	1
Quintile (nat'l) expenditure/capita	2.12	1.16	1	2	5	2.11	1.12	1	2	5
Quintile (intra-province) expenditure/capita	2.28	1.25	1	2	5	2.27	1.22	1	2	5
	Attritors ($N = 771$)									
Expenditure/capita (000s Rp)	323	272	54	252	2927					
Food expenditure/capita (000s Rp)	180	119	38	150	1073					
Non-food expenditure/capita (000s Rp)	142	197	10	86	2497					
Education expenditure/capita (000s Rp)	7	27	0	0.4	563					
Health expenditure/capita (000s Rp)	13	55	0	2	750					
Below poverty line	0.14	0.35	0	0	1					
Quintile (nat'l) expenditure/capita	3.23	1.41	1	3	5					
Quintile (intra-province) expenditure/capita	3.24	1.42	1	3	5					

Notes: See Appendix 2.8.5 for details on the panel construction. $D = d$ recipients obtained d UCT disbursements by enumeration in early 2006. Attritors are those households which could be identified in the 2005 baseline survey but not in the subsequent rounds. *Rp* stands for *Rupiah*. The exchange rate fluctuated between 9,500 Rp and 10,500 Rp to the dollar between October 2005 and September 2006. All expenditure variables are household per capita expressed in *Rupiah* per month. The underlying food expenditure items are recorded for the week prior to enumeration and scaled up to the monthly level by the factor 30/7. The underlying non-food expenditure items are recorded for the year prior to enumeration and scaled down to the monthly level by the factor 1/12. *Below poverty line* is an indicator for whether or not the household's total expenditures per capita fell below the district rural or urban poverty line in the given year. Per capita expenditure quintiles are computed separately within the full national sample and within the 31 provinces in which sample households reside. The 2005 quintiles are calculated including attritors.

Table 2.2: Propensity score model, $Pr(disbursements_h > 0|\mathbf{X}_h)$

	coefficient	**(std. error)**		**coefficient**	**(std. error)**
Urban Area	-0.138	(0.091)	*Housing status (reference = other)*		
HH Head Female	0.609	(0.110)***	Own house	-0.126	(0.102)
Land owned (hectares)	-0.112	(0.036)***	Lease house	-0.172	(0.242)
Land owned^2 (hectares)	0.001	(0.000)***	Rent house	-0.454	(0.252)*
HH ever participate Rice for the Poor	0.889	(0.075)***	Free house	-0.021	(0.207)
# children in school	-0.101	(0.063)	Official house	-0.937	(0.297)***
# children in school ^{2}	0.024	(0.017)			
Indicators for Household Size$\in \{2, \ldots, 12\}$	—	[0.222]	*Roof type (reference = other)*		
Floor area	-0.006	(0.002)***	Concrete roof	-0.697	(0.473)
			Tile roof	-0.378	(0.341)
Household composition (reference = share adult males, 10+ years)			Shingle roof	-0.121	(0.368)
Share Female Children, 0-9 yrs	0.334	(0.289)	Iron roof	-0.185	(0.303)
Share Male Children, 0-9 yrs	0.491	(0.266)*	Asbestos roof	-0.012	(0.464)
Share Adult Females, 10+ yrs	-0.090	(0.193)	Fiber/Thatch roof	0.021	(0.302)
Primary household income source (reference = other)			*Floor type (reference = other)*		
Trade/Retail	-0.155	(0.102)	Brick wall	-0.205	(0.259)
Financial/Real Estate	-0.575	(0.186)***	Wood wall	0.218	(0.282)
Agriculture	0.103	(0.075)	Bamboo wall	0.542	(0.287)*
Mining	-0.120	(0.120)	Cement/Tile/Plaster floor	-0.006	(0.473)
Manufacturing	0.198	(0.117)*	Wood/Reed/Bamboo floor	0.218	(0.484)
Electricity/Gas/Water	0.422	(0.669)	Earthen floor	0.592	(0.481)
Construction	0.307	(0.094)***			
			Souce of drinking water (reerence = other)		
Household head education level (reference = no education)			Bottled water	-0.987	(0.474)**
Primary	-0.271	(0.115)**	Pump water	-1.092	(0.252)***
Junior secondary	-0.558	(0.162)***	Tap water	-0.473	(0.295)
Senior secondary	-1.089	(0.139)***	Protected well water	-0.740	(0.249)***
Higher	-2.388	(0.511)***	Unprotected well water	-0.820	(0.265)***
			Protected spring water	-1.072	(0.280)***
Toilet facilities (reference = other)			Unprotected spring water	-1.024	(0.304)***
Own toilet	-0.254	(0.176)	River water	-0.840	(0.322)***
Shared toilet	-0.043	(0.166)	Rain water	-0.462	(0.376)
Public toilet	-0.031	(0.199)	Buy drinking water	-0.151	(0.179)
Source of light (reference = other)			*Toilet disposal location (reference = other)*		
PLN electricity	-0.450	(0.642)	Septic tank	-0.321	(0.150)**
Non-PLN electricity	-0.681	(0.763)	Pond/Rice field	-0.114	(0.225)
Pump lantern	0.352	(0.702)	Lake, river, sea	-0.106	(0.169)
Oil lamp	0.028	(0.639)	Beach	-0.105	(0.149)
Constant	1.216	(0.895)			
Pseudo-R^2	0.23				

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; Estimated using balanced panel containing 9,048 households from the *Susenas* 2005 and 2006 panel. Standard errors are clustered by village. All variables are as reported in January-March 2005. The regression also controls for province fixed effects. *PLN* is the state-run electricity firm.

Table 2.3: Idiosyncratic vs. Spatial Variation in Staggering

Fixed Effects	Province	District	Subdistrict	Province	District	Subdistrict
$\mathbf{X}_{h,t-1}$ controls	No	No	No	Yes	Yes	Yes
Dependent variable	$Pr(disbursements > 0)$					
	(1)	(2)	(3)	(4)	(5)	(6)
$H_0: \boldsymbol{\beta}_X = \mathbf{0}$						
F statistic	—	—	—	31.35	29.03	28.69
[p-value]	—	—	—	[< 0.001]	[< 0.001]	[< 0.001]
R^2	0.053	0.171	0.240	0.241	0.331	0.389
Dependent variable	$Pr(disbursements = 2 \mid disbursements > 0)$					
	(7)	(8)	(9)	(10)	(11)	(12)
$H_0: \boldsymbol{\beta}_X = \mathbf{0}$						
F statistic	—	—	—	3.01	1.81	0.90
[p-value]	—	—	—	[< 0.001]	[< 0.001]	[0.706]
R^2	0.266	0.810	0.895	0.327	0.821	0.898

Notes: *disbursements* denotes the number of disbursements received by *Susenas* enumeration in early 2006. Linear probability regressions for $Pr(disbursements_{hv} = 2 \mid disbursements_{hv} > 0)$ and $Pr(disbursements > 0)$ are based on the sample of recipient and all households, respectively. The F tests correspond to a test of the null the hypothesis that all household-specific variables included in the $\mathbf{X}$ vector have no relationship with these probabilities. The R^2 are inclusive of the geographic fixed effects. There are 30 provinces, 339 districts, and 619 subdistricts.

Table 2.4: Staggering is Orthogonal to Observable Differences Across Regions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
days since Feb. 2	-0.002	-0.001	-0.001	-0.001	-0.001	-0.001	-0.001	-0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
log distance to subdistrict capital		0.057	0.048	0.043	0.043	0.045	0.038	0.035
		(0.039)	(0.040)	(0.041)	(0.041)	(0.041)	(0.041)	(0.041)
log distance to district capital		-0.040	-0.030	-0.024	-0.026	-0.026	-0.021	-0.020
		(0.037)	(0.037)	(0.037)	(0.037)	(0.037)	(0.034)	(0.033)
log distance to Jakarta		0.018	-0.024	-0.021	-0.024	-0.019	-0.019	-0.009
		(0.032)	(0.041)	(0.040)	(0.040)	(0.044)	(0.043)	(0.040)
urban village		0.038	0.053	0.031	0.037	0.041	0.029	0.037
		(0.052)	(0.051)	(0.054)	(0.058)	(0.058)	(0.052)	(0.051)
village road paved		0.029	0.024	0.009	0.012	0.013	0.022	0.023
		(0.056)	(0.055)	(0.053)	(0.053)	(0.053)	(0.051)	(0.050)
village accessible only by water		-0.101	-0.103	-0.104	-0.101	-0.106	-0.117	-0.131
		(0.083)	(0.083)	(0.080)	(0.080)	(0.080)	(0.077)	(0.075)*
post office in village		-0.025	-0.065	-0.065	-0.064	-0.061	-0.060	-0.073
		(0.086)	(0.091)	(0.094)	(0.093)	(0.091)	(0.080)	(0.081)
log distance to post office		-0.016	-0.031	-0.009	-0.011	-0.010	-0.023	-0.030
		(0.034)	(0.035)	(0.037)	(0.037)	(0.037)	(0.034)	(0.035)
log district population, 2005			-0.083	-0.087	-0.086	-0.090	-0.097	-0.108
			(0.047)*	(0.046)*	(0.046)*	(0.047)*	(0.046)**	(0.045)**
bank in subdistrict				0.101	0.101	0.098	0.076	0.062
				(0.058)*	(0.057)*	(0.058)*	(0.057)	(0.055)
log mean HH exp./capita in district, 2005					-0.044	-0.049	-0.057	-0.058
					(0.087)	(0.088)	(0.090)	(0.086)
rainfall shock, 2005						0.076	0.081	0.117
						(0.168)	(0.156)	(0.154)
any mudslide, 2003-5							-0.177	-0.189
							(0.093)*	(0.093)**
any flood, 2003-5							-0.004	0.007
							(0.056)	(0.055)
any earthquake, 2003-5							-0.248	-0.240
							(0.122)**	(0.131)*
any fire 2003-5							-0.062	-0.064
							(0.079)	(0.079)
any other disaster, 2003-5							-0.001	0.018
							(0.090)	(0.093)
President's party 1st in village								0.137
								(0.090)
President's party 2nd in village								0.079
								(0.086)
President's party 3rd in village								0.075
								(0.095)
President's party 4th in village								-0.029
								(0.105)
Number of Households	2,349	2,349	2,349	2,349	2,349	2,349	2,349	2,349
R^2	0.001	0.020	0.035	0.047	0.048	0.049	0.091	0.105
$H_0: \boldsymbol{\beta}_X = \mathbf{0}$ [p-value]	0.470	0.591	0.120	0.024	0.017	0.026	0.001	0.0002

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; Linear probability regressions based on the sample of recipient households using the following specification: $Pr(disbursements_{hv} = 2 \mid disbursements_{hv} > 0) = \boldsymbol{\gamma}\mathbf{Z}_v + \upsilon_{hv}$, where $\mathbf{Z}_v$ is a vector of characteristics associated with the village v within which household h resides. Standard errors are clustered at the district level in all specifications. Distances to (sub)district capitals and post offices are based on actual travel distance; distance to Jakarta is great-circle. Further background on each of the variables can be found in Appendix 2.8.5.

Table 2.5: Baseline Estimates of Multi-valued Treatment Effects

Estimator	OLS	IPW	Double Robust		Control
			$(\widehat{P}_h)$	$(\mathbf{X}_h)$	Function
	(1)	(2)	(3)	(4)	(5)
			Short-Term: 2005-2006		
τ_{10}: disbursement 1	-0.062	-0.091	-0.091	-0.090	-0.076
	$(0.027)^{**}$	$(0.032)^{***}$	$(0.030)^{***}$	$(0.031)^{***}$	$(0.030)^{**}$
τ_{21}: disbursement 2	0.049	0.074	0.077	0.072	0.076
	$(0.030)^{*}$	$(0.035)^{**}$	$(0.032)^{**}$	$(0.032)^{**}$	$(0.034)^{**}$
$\tau_{20} \equiv \tau_{21} + \tau_{10}$: disbursements 1+2	-0.013	-0.017	-0.014	-0.018	0.000
	(0.014)	(0.020)	(0.020)	(0.018)	(0.019)
Reweighted	No	Yes	Yes	Yes	Yes
Propensity Score Control(s)	No	No	Yes	No	Yes
$\mathbf{X}_h$ Controls	No	No	No	Yes	No
Number of Households	9,011	9,011	9,011	9,011	9,011
R^2	0.045	0.100	0.104	0.181	0.119
			Medium-Term: 2005-2007		
τ_{10}: disbursement 1	-0.034	-0.056	-0.066	-0.044	-0.027
	(0.040)	(0.038)	$(0.038)^{*}$	(0.032)	(0.037)
τ_{21}: disbursement 2	0.026	0.028	0.032	0.009	0.028
	(0.045)	(0.043)	(0.043)	(0.035)	(0.042)
$\tau_{20} \equiv \tau_{21} + \tau_{10}$: disbursements 1+2	-0.008	-0.028	-0.034	-0.034	0.001
	(0.020)	(0.023)	(0.024)	(0.023)	(0.023)
Reweighted	No	Yes	Yes	Yes	Yes
Propensity Score Control(s)	No	No	Yes	No	Yes
$\mathbf{X}_h$ Controls	No	No	No	Yes	No
Number of Households	6,992	6,992	6,992	6,992	6,992
R^2	0.044	0.055	0.062	0.144	0.068

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The dependent variable in all specifications is Δlog total monthly household expenditures per capita between 2005 and 2006/2007. In the top panel, the constant term in columns 1 and 2 (i.e., average non-recipient log expenditure growth, or κ in equation (2.10)) equals 0.107 and 0.109, respectively. In the bottom panel, the constant term equals 0.113 and 0.153 in columns 1 and 2, respectively. Columns 2-5 are estimated by weighted least squares where the weights for treatment households equal one and the weights for control households are given by the normalized estimated odds of treatment $\widehat{\omega} = \widehat{P}/(1 - \widehat{P})$, where the normalization is over the entire sample for the given time horizon. Using the Crump et al. (2009) procedure, we trim 37 households in the upper tail of the estimated propensity scores, $\widehat{P}$, and do the same in the OLS regressions for comparability. Column 3 controls linearly for the propensity score and column 5 for a fifth-order polynomial in the propensity score allowing it to vary by treatment and control. Column 4 controls for all covariates $\mathbf{X}_h$ used to estimate the propensity score. The coefficients on these additional terms in columns 3-5 are suppressed for presentational purposes. All columns include province fixed effects. Standard errors are clustered by village and computed over the entire two-step process using a block bootstrap with 1000 repetitions.

Table 2.6: Heterogeneity by Time of Survey and Length of Delay

Estimator	OLS	IPW	Double Robust		Control
			($\widehat{P}_h$)	($\mathbf{X}_h$)	Function
	(1)	(2)	(3)	(4)	(5)
enumerated in March	0.043	0.045	0.050	0.038	0.053
	(0.024)*	(0.041)	(0.040)	(0.035)	(0.041)
τ_{10}^{f}: disbursement 1	0.003	-0.042	-0.041	-0.026	-0.029
	(0.058)	(0.064)	(0.061)	(0.063)	(0.062)
τ_{10}^{m}: disbursement 1 $\times$ enumerated in March	-0.004	0.046	0.052	0.014	0.056
	(0.064)	(0.071)	(0.070)	(0.067)	(0.070)
τ_{21}^{f}: disbursement 2	-0.085	-0.064	-0.065	-0.083	-0.059
	(0.066)	(0.072)	(0.071)	(0.071)	(0.070)
τ_{21}^{m}: disbursement 2 $\times$ enumerated in March	0.071	0.037	0.033	0.076	0.027
	(0.072)	(0.076)	(0.075)	(0.075)	(0.074)
$\tau_{10}^{March} \equiv \tau_{10}^{f} + \tau_{10}^{m}$	-0.082	-0.105	-0.106	-0.109	-0.088
	(0.030)***	(0.037)***	(0.036)***	(0.033)***	(0.034)***
$\tau_{21}^{March} \equiv \tau_{21}^{f} + \tau_{21}^{m}$	0.067	0.083	0.085	0.090	0.083
	0.033)**	(0.037)**	(0.037)**	(0.034)***	(0.036)**
$\tau_{20}^{March} \equiv \tau_{21}^{f} + \tau_{10}^{f} + \tau_{21}^{m} + \tau_{10}^{m}$	-0.015	-0.022	-0.021	-0.019	-0.005
	(0.015)	(0.021)	(0.020)	(0.019)	(0.020)
$\tau_{20}^{February} \equiv \tau_{10}^{f} + \tau_{21}^{f}$	-0.001	0.004	0.011	-0.012	0.027
	(0.032)	(0.049)	(0.048)	(0.043)	(0.050)
Reweighted	No	Yes	Yes	Yes	Yes
Propensity Score Control(s)	No	No	Yes	No	Yes
$\mathbf{X}_h$ Controls	No	No	No	Yes	No
Number of Households	9,011	9,011	9,011	9,011	9,011

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; Significance levels: $*$: 10% $**$: 5% $***$: 1%; The dependent variable in all specifications is Δlog household expenditures per capita between 2005 and 2006. The variable *enumerated in March* is an indicator for whether the household was enumerated in March (relative to February) 2006. All columns are estimated by weighted least squares where the weights for treatment households equal one and the weights for control households are given by the normalized estimated odds of treatment $\widehat{\omega} = \widehat{P}/(1 - \widehat{P})$. See Table 2.5 for further details on each of the estimators. All columns include province fixed effects. Standard errors are clustered by village and computed using a block bootstrap with 1000 repetitions.

Table 2.7: Multi-valued Treatment Effects by Expenditure Group

Growth Horizon	2005-2006			2005-2007		
Expenditure Group	**total**	**food**	**non-food**	**total**	**food**	**non-food**
	(1)	(2)	(3)	(4)	(5)	(6)
τ_{10}: disbursement 1	-0.076	-0.094	-0.046	-0.017	-0.030	0.031
	(0.031)**	(0.031)***	(0.045)	(0.038)	(0.039)	(0.056)
τ_{21}: disbursement 2	0.076	0.097	0.037	0.023	0.068	-0.053
	(0.034)**	(0.034)***	(0.048)	(0.043)	(0.041)*	(0.060)
$\tau_{20} \equiv \tau_{21} + \tau_{10}$: disbursements 1+2	0.000	0.003	-0.010	0.006	0.038	-0.022
	(0.019)	(0.019)	(0.025)	(0.026)	(0.026)	(0.039)
Number of Households	9,011	9,011	9,009	6,992	6,992	6,992

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The dependent variable in all specifications is Δlog household expenditures per capita on the given commodity group between 2005 and 2006/2007. All columns are estimated by weighted least squares where the weights for treatment households equal one and the weights for control households are given by the normalized estimated odds of treatment $\widehat{\omega} = \widehat{P}/(1-\widehat{P})$. All columns include a 5th order polynomial in the propensity scores that is allowed to vary by treatment and control. All columns include province fixed effects. Standard errors are clustered by village and computed using a block bootstrap with 1000 repetitions.

Table 2.8: Multi-valued Treatment Effects by Expenditure Subgroup, 2005-2006

Expenditure Group	**grains** (1)	**fish/meat/dairy** (2)	**fruit/nuts/veg.** (3)	**other food** (4)	**outside prep.** (5)
τ_{10}: disbursement 1	-0.025	-0.173	-0.120	-0.084	-0.218
	(0.045)	(0.063)***	(0.043)***	(0.040)**	(0.071)***
τ_{21}: disbursement 2	0.055	0.125	0.122	0.085	0.160
	(0.054)	(0.068)*	(0.051)**	(0.045)*	(0.079)**
$\tau_{20} \equiv \tau_{21} + \tau_{10}$: disbursements 1+2	0.031	-0.048	0.003	0.000	-0.059
	(0.034)	(0.035)	(0.030)	(0.023)	(0.041)
Number of Households	8,789	8,334	8,844	8,879	7,649
Expenditure Group	**housing** (6)	**transport/comm.** (7)	**appliances** (8)	**debt/taxes** (9)	**educ./health** (10)
τ_{10}: disbursement 1	0.012	-0.285	0.006	0.083	-0.149
	(0.052)	(0.118)**	(0.065)	(0.122)	(0.109)
τ_{21}: disbursement 2	-0.031	0.259	0.119	-0.149	0.191
	(0.056)	(0.132)**	(0.074)	(0.135)	(0.125)
$\tau_{20} \equiv \tau_{21} + \tau_{10}$: disbursements 1+2	-0.020	-0.026	0.125	-0.066	0.042
	(0.025)	(0.061)	(0.046)***	(0.059)	(0.065)
Number of Households	9,002	5,478	8,898	5,995	6,504

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The dependent variable in all specifications is Δlog household expenditures per capita on the given commodity group between 2005 and 2006/2007. All columns are estimated by weighted least squares where the weights for treatment households equal one and the weights for control households are given by the normalized estimated odds of treatment $\widehat{\omega} = \widehat{P}/(1 - \widehat{P})$. All columns include a 5th order polynomial in the propensity scores that is allowed to vary by treatment and control. All columns include province fixed effects. Standard errors are clustered by village and computed using a block bootstrap with 1000 repetitions.

Table 2.9: Heterogeneity by Proximity to Financial Institutions

Growth Horizon	2005-2006			2005-2007		
Estimator	OLS (1)	IPW (2)	Control Fn. (3)	OLS (4)	IPW (5)	Control Fn. (6)
bank nearby	0.001	-0.027	-0.013	0.003	0.022	0.049
	(0.026)	(0.047)	(0.045)	(0.033)	(0.044)	(0.042)
τ^u_{10}: disbursement 1	-0.133	-0.165	-0.140	-0.142	-0.144	-0.100
	(0.047)***	(0.061)***	(0.054)**	(0.060)**	(0.063)**	(0.058)*
τ^b_{10}: disbursement 1 × bank nearby	0.104	0.108	0.104	0.153	0.126	0.112
	(0.058)*	(0.069)	(0.064)	(0.077)**	(0.079)	(0.074)
τ^u_{21}: disbursement 2	0.102	0.106	0.118	0.067	0.052	0.058
	(0.054)*	(0.057)*	(0.057)**	(0.067)	(0.071)	(0.066)
τ^b_{21}: disbursement 2 × bank nearby	-0.087	-0.065	-0.078	-0.066	-0.042	-0.051
	(0.067)	(0.072)	(0.070)	(0.086)	(0.089)	(0.085)
$\tau^{banked}_{10} \equiv \tau^u_{10} + \tau^b_{10}$	-0.027	-0.057	-0.036	0.012	-0.018	0.012
	(0.032)	(0.037)	(0.035)	(0.049)	(0.047)	(0.045)
$\tau^{banked}_{21} \equiv \tau^u_{21} + \tau^b_{21}$	0.015	0.041	0.040	0.002	0.009	0.007
	(0.035)	(0.042)	(0.039)	(0.055)	(0.054)	(0.051)
$\tau^{banked}_{20} \equiv \tau^u_{21} + \tau^u_{10} + \tau^b_{21} + \tau^b_{10}$	-0.014	-0.016	0.004	0.014	-0.009	0.019
	0.015	(0.019)	(0.020)	(0.023)	(0.026)	(0.027)
$\tau^{unbanked}_{20} \equiv \tau^u_{10} + \tau^u_{21}$	-0.030	-0.058	-0.023	-0.074	-0.092	-0.042
	0.030	(0.042)	(0.039)	(0.043)*	(0.050)*	(0.049)
Reweighted	No	Yes	Yes	No	Yes	Yes
Propensity Score Controls	No	No	Yes	No	No	Yes
Number of Households	8,923	8,923	8,923	6,966	6,966	6,966

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The dependent variable in all specifications is Δlog household expenditures per capita between 2005 and 2006/2007. The variable, *bank nearby*, equals one if there are any banking institutions located in the given village's subdistrict as reported in *Podes* 2005 (see Appendix 2.8.5). Columns 2-3 and 5-6 are estimated by weighted least squares where the weights for treatment households equal one and the weights for control households are given by the normalized estimated odds of treatment $\widehat{\omega} = \widehat{P}/(1-\widehat{P})$. Columns 3 and 6 additionally include a 5th order polynomial in the propensity scores that is allowed to vary by treatment and control. All columns include province fixed effects. Standard errors are clustered by village and computed using a block bootstrap with 1000 repetitions.

Table 2.10: Robustness Checks

		τ_{10} disbursement 1	τ_{21} disbursement 2	$\tau_{20} \equiv \tau_{10} + \tau_{21}$ disbursements 1+2	No. of Households
1.	baseline	-0.076	0.076	0.000	9,011
		(0.031)**	(0.034)**	(0.019)	
2.	controls for day of survey enumeration	-0.070	0.071	0.001	9,011
		(0.027)**	(0.031)**	(0.017)	
3.	district fixed effects	-0.059	0.057	-0.002	9,011
		(0.024)**	(0.027)**	(0.016)	
4.	subdistrict fixed effects	-0.038	0.037	-0.002	9,011
		(0.022)*	(0.024)	(0.014)	
5.	village fixed effects	-0.039	0.038	-0.001	9,011
		(0.021)*	(0.024)	(0.015)	
6.	$\ln C_{h,t-1}$ trimmed at 1st and 99th percentile	-0.085	0.074	-0.011	8,833
		(0.031)***	(0.034)**	(0.018)	
7.	$\Delta \ln C_{ht}$ trimmed at 1st and 99th percentile	-0.052	0.050	-0.002	8,832
		(0.024)**	(0.027)*	(0.016)	
8.	$\Delta \ln C_{ht}$ deflated by regional poverty line	-0.057	0.066	0.009	8,851
		(0.033)*	(0.036)*	(0.019)	
9.	$\Delta \ln C_{ht}$ deflated by regional CPI	-0.070	0.069	-0.001	8,923
		(0.029)**	(0.032)**	(0.017)	
10.	$\Delta \ln C_{ht}$ adjusted for adult-equivalence (total)	-0.070	0.071	0.002	9,011
		(0.030)**	(0.033)**	(0.018)	
11.	$\Delta \ln C_{ht}$ adjusted for adult-equivalence (food)	-0.072	0.072	0.001	9,011
		(0.028)**	(0.032)**	(0.018)	
12.	$\Delta \ln C_{ht}$ prorated annual durable expenditures	-0.092	0.083	-0.009	9,011
		(0.034)***	(0.038)**	(0.020)	
13.	controls for other social program receipt	-0.056	0.073	0.017	9,011
		(0.031)*	(0.032)**	(0.025)	
14.	controls for change in health shocks/capita	-0.078	0.078	0.001	9,011
		(0.030)***	(0.033)**	(0.018)	
15.	controls for natural disasters, 2003-5	-0.084	0.080	-0.003	8,785
		(0.031)***	(0.033)**	(0.018)	
16.	controls for staggering covariates, Table 2.4	-0.074	0.051	-0.024	8,687
		(0.029)**	(0.032)	(0.029)	
17.	controls for pre-program enumerator visit	-0.045	0.075	0.030	9,011
		(0.038)	(0.032)**	(0.029)	

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; Each row corresponds to a separate regression with log household expenditure per capita growth, $\Delta \ln C$, as the dependent variable. The estimates are obtained by weighted least squares where the weights for treatment households equal one and the weights for control households are given by the normalized estimated odds of treatment $\widehat{\omega} = \widehat{P}/(1-\widehat{P})$. All rows include a 5th order polynomial in the propensity scores that is allowed to vary by treatment and control. All rows include province fixed effects. Standard errors are clustered by village and computed using a block bootstrap with 1000 repetitions.

Table 2.11: Intensive Margin Treatment Effects by Expenditure Group

	(1)	(2)	(3)	(4)
	$\Delta \ln total\ expenditures/capita$			
transfers per capita (000,000s Rupiah)	0.045	0.044	0.037	0.064
	(0.008)***	(0.008)***	(0.009)***	(0.012)***
	$\Delta \ln food\ expenditures/capita$			
transfers per capita (000,000s Rupiah)	0.045	0.044	0.039	0.063
	(0.009)***	(0.009)***	(0.010)***	(0.013)***
	$\Delta \ln non-food\ expenditures/capita$			
transfers per capita (000,000s Rupiah)	0.046	0.045	0.036	0.070
	(0.011)***	(0.012)***	(0.012)***	(0.019)***
Treatment Indicators	Yes	Yes	Yes	Yes
Reweighted	No	Yes	Yes	Yes
Propensity Score Polynomial	No	No	Yes	Yes
Household Size Indicators	No	No	No	Yes
Number of Households	9,011	9,011	9,011	9,011

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; Each cell corresponds to a separate regression. Transfers are rescaled to 100,000s of Rupiah (approximately 10 USD). Columns 2-4 are estimated by weighted least squares where the weights for treatment households equal one and the weights for control households are given by the normalized estimated odds of treatment $\widehat{\omega} = \widehat{P}/(1-\widehat{P})$. All include the treatment indicators and province fixed effects. Columns 3-4 include a 5th order polynomial in the propensity scores that is allowed to vary by treatment and control. Column 4 includes indicators for household size. Standard errors are clustered by village and computed using a block bootstrap with 1000 repetitions.

Table 2.12: Household Expenditures and Transitory Rainfall Shocks

	(1)	(2)	(3)	(4)	(5)
rainfall shock	0.030	-0.049	-0.043	-0.055	-0.029
	(0.071)	(0.095)	(0.073)	(0.074)	(0.068)
rural village		-0.005			
		(0.019)			
rural village × rainfall shock		0.125			
		(0.109)			
agriculture primary income			0.018		
			(0.014)		
agricultural primary income × rainfall shock			0.220		
			(0.088)**		
own any agricultural land				-0.001	
				(0.014)	
own any agricultural land × rainfall shock				0.170	
				(0.087)*	
own agricultural land (Ha)					0.001
					(0.003)
own agricultural land (Ha) × rainfall shock					0.063
					(0.029)**
rainfall shock + rainfall shock × covariate		0.077	0.177	0.114	
		(0.084)	(0.095)*	(0.091)	
Number of Households	8,923	8,923	8,923	8,923	8,923
R^2	0.042	0.043	0.045	0.044	0.045

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The dependent variable in all specifications is Δlog total monthly household expenditures per capita between 2005 and 2006. The rainfall shock is the log deviation of the seasonal rainfall level in the district from the long-run (1952-2004) district mean. Standard errors clustered by district. All columns include province fixed effects. The interaction terms are as observed at baseline.

2.8 Appendix

2.8.1 Propensity Scores and Reconstructed Quasi-PMT Scores

To estimate the probability that household h receives any UCT disbursements, D, we consider the following specification, which roughly approximates information on household h available to enumerators and local officials in mid-2005,

$$Pr(disbursements_h > 0) = F\left(\boldsymbol{\beta}\mathbf{X}_h^{\text{fam}} + \boldsymbol{\gamma}\mathbf{X}_h^{\text{house}} + \boldsymbol{\alpha}\mathbf{X}_h^{\text{head}} + \boldsymbol{\delta}\mathbf{X}_h^{\text{welfare}} + \zeta_h > 0\right), \tag{2.12}$$

All right-hand variables are observed in at baseline in early 2005: $\mathbf{X}_h^{\text{fam}}$ is a vector of demographic variables including household age structure, gender breakdown; $\mathbf{X}_h^{\text{house}}$ includes variables pertaining to the quality of the physical structures in which household h lives; $\mathbf{X}_h^{\text{head}}$ includes characteristics of the head of the household, $\mathbf{X}_h^{\text{welfare}}$ includes indicators for employment among household members, prior participation in government welfare programs, and amount of land owned; F is the logistic CDF;[37] and ζ_h captures all variables unobservable to the econometrician but possibly observable to program administrators. We also control for province fixed effects to subsume some of the regional differences in targeting infrastructure (among other things). A full elaboration of the coefficient estimates was reported in Table 2.2.[38] Given our large set of dummy variables, there is little advantage to estimating equation (2.12) nonparametrically.[39]

As discussed in Section 2.3.2, although we made every effort to reconstruct the underlying PMT scores using available data, the resulting scores were not discriminating enough to allow for a fuzzy regression-discontinuity research design.

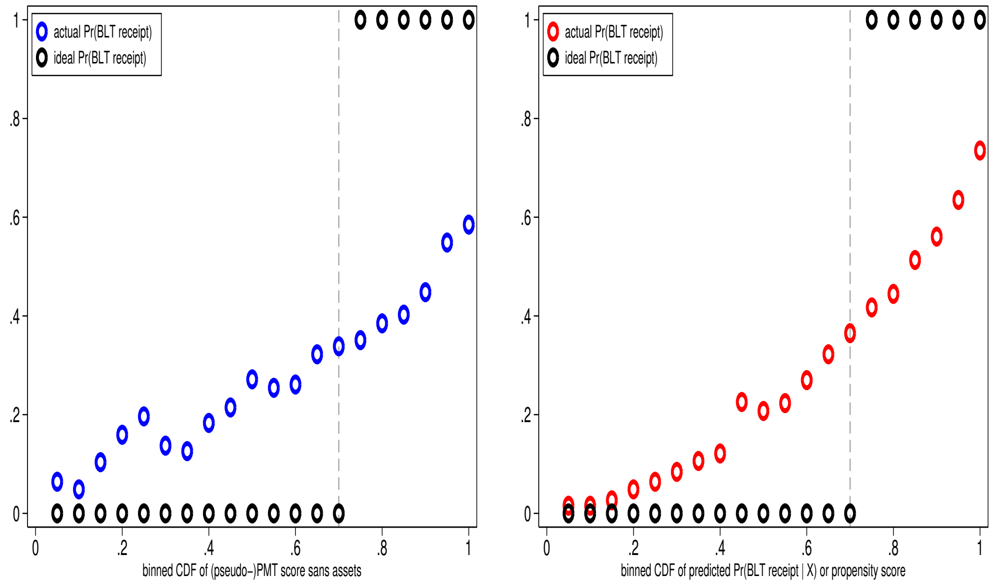
[37]In unreported results, we find that a probit estimator yields identical results.

[38]The official eligibility survey grouped several response categories to questions in *Susenas* concerning household characteristics. Whether one leaves the individual responses as separate indicators (in a fully saturated sense) or groups them according to the rubric in the original survey does not matter for the key qualitative findings in this paper.

[39]Doing so using the Klein and Spady (1993) estimator yields an estimated propensity score that has a 0.97 correlation with the simpler parametric logit.

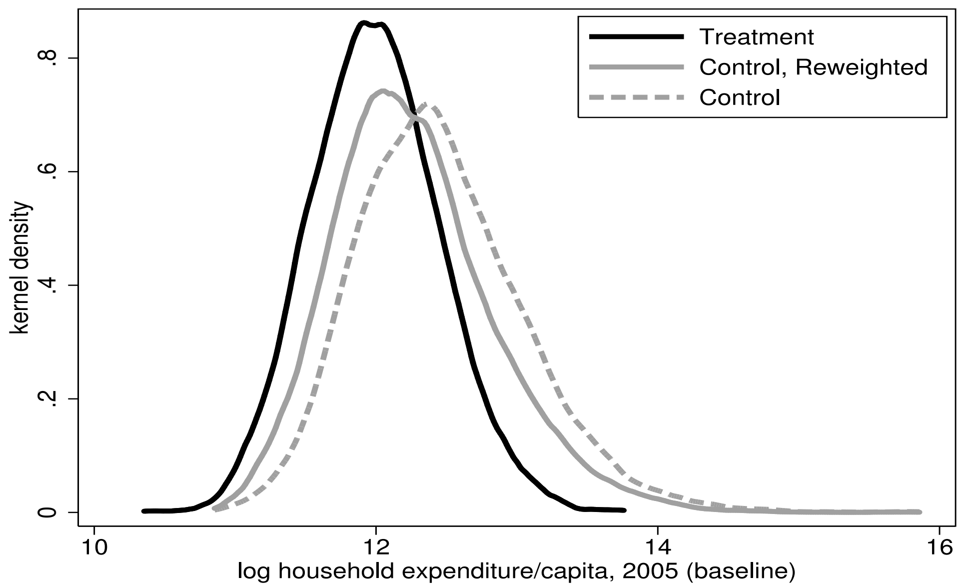
After transforming applicable questions in *Susenas* 2005 into the corresponding variable-specific eligibility criteria, we apply the district-specific PMT coefficients corresponding to the given variables to produce a measure $\widetilde{P}_h$, which reflects a data-constrained approximation to the actual PMT scores based on the original eligibility survey.[40] According to program guidelines, households with PMT scores above the 70th percentile should qualify for benefits. We take this rule to our estimates $\widetilde{P}_h$ in search of a potential discontinuity. Unfortunately, as seen in Figure 2.7, no such discontinuity can be found—perhaps unsurprisingly given the evidence on leakage and undercoverage. Moreover, the actual probability of UCT receipt looks quite similar across the distribution of the estimated propensity scores $\widehat{P}_h$. Yet, if we predict the probability of program receipt using $\widetilde{P}_h$ as the only regressor—effectively fixing $(\beta, \gamma, \delta, \alpha)$ in equation 2.12 at the district-specific PMT coefficients—and accordingly reweight households in the control group, the balance at baseline is much worse than when using our arguably more flexible approach in equation (2.12). This can be seen by comparing the effect of reweighting the control group in Figure 2.8, which uses $\widetilde{P}_h$, and Figure 2.4 discussed in the paper, which uses our estimated propensity scores. This balance differential is intuitive because our propensity score model is based on a richer set of variables plausibly in the information set of local officials engaged in community-based alongside or possibly in defiance of official targeting.

[40]Prior to this, we rescale the coefficients to ensure that they sum to 1 after dropping the questions not available in *Susenas*.



Notes: LEFT—The circles capture the share of UCT (BLT) recipients within the given bin where the bins are 0.05 width slices of the CDF of the quasi-PMT scores approximated using the procedures described in the text. The dashed vertical line constitutes the 30% threshold above which households were (in theory) supposed to receive the program. RIGHT—The circles capture the share of UCT (BLT) recipients within the given bin where the bins are 0.05 width slices of the CDF of the propensity scores obtained from estimating a binary version of equation (2.12) by maximum likelihood where ζ_h is logistic distributed.

Figure 2.7: Propensity Score Estimates and Approximated PMT Scores



Notes: All distributions estimated using Epanechnikov kernel and a rule-of-thumb bandwidth. The "Control (Reweighted)" observations are adjusted using inverse probability weights (IPW) based on normalized estimated odds of treatment $\widetilde{\omega} = \widetilde{P}/(1 - \widetilde{P})$, where $\widetilde{P}_h$ is as described in Appendix 2.8.1.

Figure 2.8: Baseline Expenditure Distributions by Treatment Status

2.8.2 Further Empirical Results

This appendix provides tables and discussion of additional empirical results mentioned in the paper.

Validating the Exogeneity of Variation in Transfers per Capita

Table 2.13 shows that the UCT program had no effect on household size over the period from early 2005 to early 2006, and Table 2.14 shows that the size of the informal tax levied on UCT recipients is uncorrelated with household characteristics. The results in these tables support the claim in the paper that the residual variation in transfers per capita is exogenous with respect to a number of other factors potentially associated with expenditure growth over the period under study.

Lack of Heterogeneity across Quantiles of Expenditure Growth

One concern with the baseline estimates of average multivalued treatment effects on the treated is that they miss important heterogeneous effects beyond the mean. We address this concern by applying the estimator in Firpo (2007) to recover quantile treatment effects on the treated (QTT) at every quantile $q \in \{5, 6, \ldots, 95\}$ of the main aggregate expenditure growth outcome from early 2005 to early 2006. The results of this exercise are reported in Figure 2.9, which compares the baseline estimates of the average treatment effects on the treated (ATT) parameters $\boldsymbol{\tau}$ in column 3 of Table 2.5 (dashed horizontal line) to the relevant QTT estimate and its associated 95% confidence interval across q.[41] The message from the three graphs is clear. The baseline ATT estimates in the paper do not obscure heterogeneous effects at other moments of the potential expenditure growth outcomes.

[41] These QTT estimates are a straightforward adaption of the reweighting (plus staggering) approach to recovering the multivalued ATT in equation (2.10) to the binary QTT approach developed in Firpo (2007). However, the baseline IPW reweighting estimator is most comparable to this approach as double robustness properties of the QTT approach have not been explored in the literature.

Heterogeneity by Urban vs. Rural

In Table 2.15, we allow the estimated treatment effects $\boldsymbol{\tau}$ to vary across urban and rural areas. In doing so, we find that household expenditures respond similarly to the UCT program in both locations.

Heterogeneity by Age of Household Head

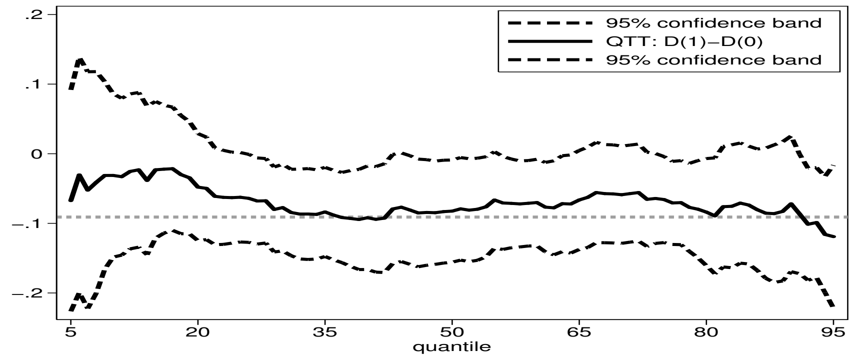
In Table 2.16, we allow the estimated treatment effects $\boldsymbol{\tau}$ to vary by the age of the household head. We do this in columns 1 and 3 by interacting the treatment indicators with a quadratic in age and in columns 2 and 4 by interacting those indicators with three bins comprising roughly the bottom three quintiles of the age distribution. The marginal effects reported in Figure 2.6 correspond to the quadratic interactions in columns 1 and 3. While we observe an interesting nonlinear age profile in the treatment effects consistent with life cycle behavior, the estimates underlying Figure 2.6 are somewhat imprecise.

Heterogeneous Expenditure Response to Transitory Rainfall Shocks

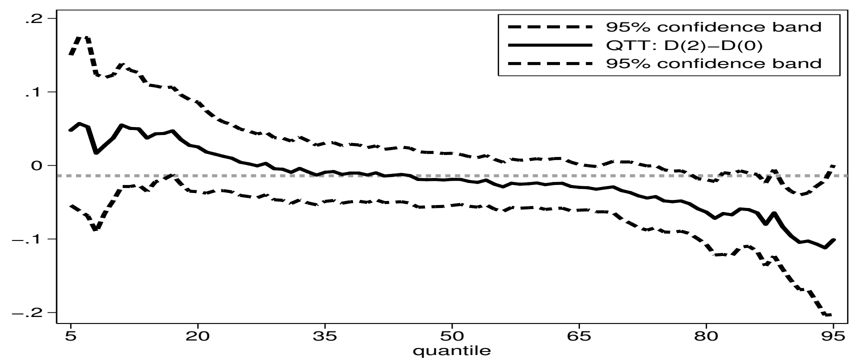
In Table 2.12 of the paper, we showed that household expenditures were excessively sensitive to transitory rainfall shocks for certain segments of the (rural) population engaged in agricultural activity. In Table 2.17 below, we show that these patterns cannot arise spuriously. More specifically, we find that the UCT benefits do not have heterogeneous effects along the same dimensions of exposure to productivity shocks including rural residence, agriculture being the main source of household income, owning any land, amount of land owned. These results further substantiate the role of shock-specific assets in mediating the expenditure response to transitory income shocks. In other words, the transitory income shock implicit in the UCT program had similar effects across households because the returns to the UCT did not depend on specific complementary assets.

2.8.3 Figures

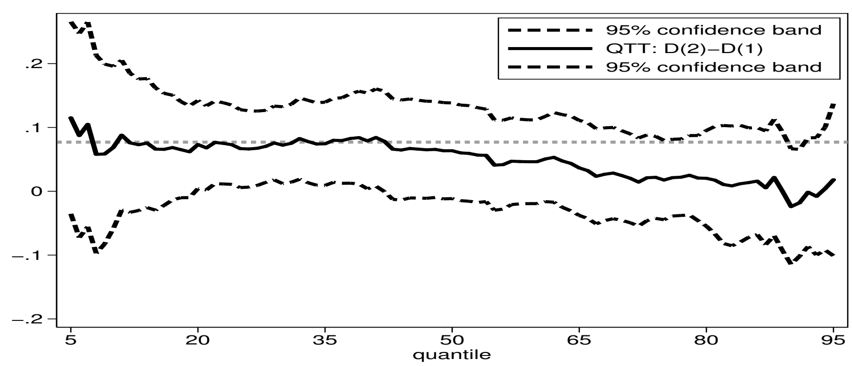
QTT_{10}: 1 relative to 0 disbursements



QTT_{20}: 2 relative to 0 disbursements



QTT_{21}: 2 relative to 1 disbursements



Notes: These figures report the results of estimating the quantile treatment effects on the treated (QTT) analogue to equation (2.10) based on a straightforward adaption of the approach developed in Firpo (2007). The dashed horizontal line corresponds to the ATT estimates of $\boldsymbol{\tau}$ reported in column 2 of Table 2.5. The 95% confidence bands are based on 300 block-bootstrap repetitions, estimating the QTT parameters across all q during each repetition.

Figure 2.9: Quantile Treatment Effects on the Treated (QTT)

2.8.4 Tables

Table 2.13: UCT Benefits Had No Effect on Household Size

Estimator	OLS	IPW	Double Robust		Control
			($\widehat{P}_h$)	($\mathbf{X}_h$)	Function
	(1)	(2)	(3)	(4)	(5)
τ_{10}: disbursement 1	-0.014	0.069	0.069	0.058	0.095
	(0.059)	(0.127)	(0.129)	(0.116)	(0.117)
τ_{21}: disbursement 2	0.014	-0.090	-0.090	-0.045	-0.115
	(0.065)	(0.133)	(0.135)	(0.116)	(0.127)
$\tau_{20} \equiv \tau_{21} + \tau_{10}$: disbursements 1+2	0.0002	-0.021	-0.022	0.013	-0.020
	(0.030)	(0.078)	(0.079)	(0.075)	(0.080)
Reweighted	No	Yes	Yes	Yes	Yes
Propensity Score Control(s)	No	No	Yes	No	Yes
$\mathbf{X}_h$ Controls	No	No	No	Yes	No
Number of Households	9,011	9,011	9,011	9,011	9,011
R^2	0.006	0.011	0.011	0.131	0.016

Notes: All columns estimated by linear probability regressions with Δ log household size between 2005 and 2006 on the left hand side. Columns 2-5 are estimated by weighted least squares where the weights for treatment households equal one and the weights for control households are given by the normalized $\widehat{\omega} = \widehat{P}/(1 - \widehat{P})$. All columns include province fixed effects. Standard errors are clustered by village and computed using a block bootstrap with 1000 repetitions.

Table 2.14: Idiosyncratic vs. Spatial Variation in the "Tax" on UCT Recipients

	(1)	(2)	(3)	(4)
household size, $t-1$	-0.001	0.001	0.003	-0.000
	(0.004)	(0.002)	(0.005)	(0.002)
Fixed Effects (FE)	Province	Subdistrict	Province	Subdistrict
$\mathbf{X}_{h,t-1}$ controls	No	No	Yes	Yes
p-value joint statistical significance			[0.52]	[0.99]
Number of Households	2,410	2,410	2,410	2,410
R^2	0.113	0.822	0.187	0.827

Notes: All columns are estimated by linear probability regressions of the following specification: $Pr(transfer_h < full\ amount \mid disbursements > 0) = \boldsymbol{\beta}\mathbf{X}_{h,t-1} + \theta_{FE} + e_h$, where $\mathbf{X}_{h,t-1}$ includes all the baseline household characteristics used to estimate propensity scores. Standard errors clustered by village.

Table 2.15: (Lack of) Heterogeneity by Urban vs. Rural Location

Growth Horizon	2005-2006			2005-2007		
Expenditure Group	**total** (1)	**food** (2)	**non-food** (3)	**total** (4)	**food** (5)	**non-food** (6)
rural	-0.006	-0.029	0.013	-0.103	-0.065	-0.136
	(0.031)	(0.031)	(0.040)	(0.036)***	(0.033)**	(0.054)**
τ_{10}^{u}: disbursement 1	-0.086	-0.116	-0.022	0.067	0.053	0.151
	(0.059)	(0.058)**	(0.078)	(0.076)	(0.080)	(0.097)
τ_{10}^{r}: disbursement 1 × rural	0.015	0.038	-0.036	-0.080	-0.085	-0.123
	(0.066)	(0.063)	(0.095)	(0.087)	(0.085)	(0.113)
τ_{21}^{u}: disbursement 2	0.089	0.109	0.033	0.070	0.088	-0.015
	(0.064)	(0.065)*	(0.084)	(0.084)	(0.083)	(0.113)
τ_{21}^{r}: disbursement 2 × rural	-0.017	-0.016	0.004	-0.066	-0.035	-0.050
	(0.069)	(0.071)	(0.100)	(0.095)	(0.090)	(0.126)
$\tau_{10}^{rural} \equiv \tau_{10}^{u} + \tau_{10}^{r}$:	-0.071	-0.078	-0.058	-0.013	-0.032	0.027
	(0.034)**	(0.035)**	(0.056)	(0.043)	(0.039)	(0.060)
$\tau_{21}^{rural} \equiv \tau_{21}^{u} + \tau_{21}^{r}$	0.072	0.093	0.037	0.004	0.053	-0.064
	(0.038)*	(0.038)**	(0.058)	(0.047)	(0.045)	(0.063)
$\tau_{20}^{rural} \equiv \tau_{21}^{u} + \tau_{10}^{u} + \tau_{21}^{r} + \tau_{10}^{r}$	0.001	0.015	-0.021	-0.009	0.021	-0.037
	(0.023)	(0.023)	(0.034)	(0.026)	(0.027)	(0.041)
$\tau_{20}^{urban} \equiv \tau_{10}^{u} + \tau_{21}^{u}$	0.003	-0.007	0.011	0.137	0.141	0.136
	(0.026)	(0.026)	(0.036)	(0.046)***	(0.038)***	(0.064)**
Number of Households	9,011	9,011	9,009	6,992	6,992	6,992

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The dependent variable in all specifications is Δlog household expenditures per capita on the given commodity group between 2005 and 2006/2007. All columns are estimated by weighted least squares where the weights for treatment households equal one and the weights for control households are given by the normalized $\widehat{\omega} = \widehat{P}/(1-\widehat{P})$. The measure of rural location is based on a designation assigned by the Central Bureau of Statistics (BPS) to every village in the country based on population density, proximity to public health and education facilities, and existence of agricultural activities. All columns include a 5th order polynomial in the propensity scores that is allowed to vary by treatment and control. All columns include province fixed effects. Standard errors are clustered by village and computed using a block bootstrap with 1000 repetitions.

Table 2.16: Heterogeneity by the Age of the Household Head

Growth Horizon	2005-2006		2005-2007	
	(1)	(2)	(3)	(4)
disbursement 1	0.235	0.014	0.336	-0.006
	(0.231)	(0.042)	(0.258)	(0.056)
disbursement 1 × age of HH head	-0.016		-0.014	
	(0.009)*		(0.011)	
disbursement 1 × age of HH head squared	0.000		0.000	
	(0.000)**		(0.000)	
disbursement 1 × age of HH head $\in [13, 36]$		-0.086		0.056
		(0.058)		(0.076)
disbursement 1 × age of HH head $\in [37, 46]$		-0.116		-0.013
		(0.055)**		(0.077)
disbursement 1 × age of HH head $\in [47, 56]$		-0.183		-0.067
		(0.062)***		(0.074)
disbursement 2	0.234	0.025	-0.189	0.060
	(0.212)	(0.039)	(0.264)	(0.053)
disbursement 2 × age of HH head	-0.005		0.007	
	(0.008)		(0.011)	
disbursement 2 × age of HH head squared	0.000		-0.000	
	(0.000)		(0.000)	
disbursement 2 × age of HH head $\in [13, 36]$		0.082		-0.081
		(0.050)		(0.069)
disbursement 2 × age of HH head $\in [37, 46]$		0.069		-0.020
		(0.048)		(0.069)
disbursement 2 × age of HH head $\in [47, 56]$		0.075		-0.067
		(0.049)		(0.067)
disbursement 1+2	0.469	0.039	0.147	0.054
	(0.191)**	(0.037)	(0.189)	(0.044)
disbursement 1+2 × age of HH head	-0.020		-0.007	
	(0.008)***		(0.008)	
disbursement 1+2 × age of HH head squared	0.000		0.000	
	(0.000)***		(0.000)	
disbursement 1+2 × age of HH head $\in [13, 36]$		-0.004		-0.025
		(0.049)		(0.062)
disbursement 1+2 × age of HH head $\in [37, 46]$		-0.047		-0.033
		(0.041)		(0.058)
disbursement 1+2 × age of HH head $\in [47, 56]$		-0.108		-0.133
		(0.050)**		(0.056)**
Number of Households	9,011	9,011	6,992	6,992

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The dependent variable in all specifications is Δlog household expenditures per capita between 2005 and 2006/2007. All columns are estimated by weighted least squares where the weights for treatment households equal one and the weights for control households are given by the normalized $\widehat{\omega} = \widehat{P}/(1 - \widehat{P})$. All columns include a 5th order polynomial in the propensity scores that is allowed to vary by treatment and control. All columns include province fixed effects. Standard errors are clustered by village and computed using a block bootstrap with 1000 repetitions.

Table 2.17: Validating the Differential Response to Transitory Rainfall Shocks

Covariate	—	rural village	agri. main income	own any agri. land	agri. land (Ha)
	(1)	(2)	(3)	(4)	(5)
rainfall shock	0.035	-0.064	-0.110	-0.249	-0.096
	(0.136)	(0.116)	(0.107)	(0.142)*	(0.119)
τ_{10}: receipt of disbursement 1	-0.068	-0.080	-0.111	-0.084	-0.059
	(0.028)**	(0.056)	(0.035)***	(0.034)**	(0.030)**
τ_{21}: receipt of disbursement 2	0.067	0.084	0.095	0.078	0.061
	(0.032)**	(0.062)	(0.040)**	(0.040)*	(0.034)*
covariate		-0.004	0.000	0.006	0.001
		(0.028)	(0.029)	(0.024)	(0.005)
covariate × rainfall shock		0.131	0.335	0.493	0.135
		(0.168)	(0.131)**	(0.150)***	(0.049)***
covariate × disbursement 1		0.013	0.075	0.006	-0.017
		(0.064)	(0.045)*	(0.050)	(0.014)
covariate × disbursement 2		-0.020	-0.052	0.002	0.015
		(0.069)	(0.048)	(0.053)	(0.015)
Number of Households	8,923	8,923	8,923	8,923	8,923

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; The dependent variable in all specifications is Δlog total monthly household expenditures per capita between 2005 and 2006. The rainfall shock is the log deviation of the seasonal rainfall level in the district from the long-run (1952-2004) district mean. Standard errors clustered by district. All columns include province fixed effects. The interaction terms are as observed at baseline.

Table 2.18: Reassigning Control Households to Treatment Groups

Estimator	OLS	IPW	Double Robust		Control
			$(\widehat{P}_h)$	$(\mathbf{X}_h)$	Function
	(1)	(2)	(3)	(4)	(5)
			Short-Term: 2005-2006		
τ_{10}: disbursement 1	-0.051	-0.077	-0.076	-0.081	-0.047
	(0.030)*	(0.035)**	(0.033)**	(0.033)**	(0.032)
τ_{21}: disbursement 2	0.022	0.036	0.038	0.042	0.034
	(0.032)	(0.036)	(0.036)	(0.033)	(0.035)
$\tau_{20} \equiv \tau_{21} + \tau_{10}$: disbursements 1+2	-0.030	-0.040	-0.038	-0.039	-0.013
	(0.015)**	(0.020)**	(0.020)*	(0.019)**	(0.018)
Reweighted	No	Yes	Yes	Yes	Yes
Propensity Score Control(s)	No	No	Yes	No	Yes
$\mathbf{X}_h$ Controls	No	No	No	Yes	No
Number of Households	9,007	9,007	9,007	9,007	9,007
R^2	0.045	0.079	0.083	0.153	0.099
			Medium-Term: 2005-2007		
τ_{10}: disbursement 1	0.005	-0.012	-0.016	-0.003	0.016
	(0.040)	(0.039)	(0.039)	(0.036)	(0.040)
τ_{21}: disbursement 2	-0.019	-0.016	-0.015	-0.022	-0.020
	(0.043)	(0.043)	(0.041)	(0.037)	(0.043)
$\tau_{20} \equiv \tau_{21} + \tau_{10}$: disbursements 1+2	-0.014	-0.028	-0.031	-0.026	-0.005
	(0.020)	(0.022)	(0.024)	(0.022)	(0.022)
Reweighted	No	Yes	Yes	Yes	Yes
Propensity Score Control(s)	No	No	Yes	No	Yes
$\mathbf{X}_h$ Controls	No	No	No	Yes	No
Number of Households	6,994	6,994	6,994	6,994	6,994
R^2	0.044	0.059	0.062	0.144	0.074

Notes: Significance levels: $*$: 10% $**$: 5% $***$: 1%; This table reproduces the baseline estimates after first reassigning all control households ($D = 0$) to one of the two treatment levels ($D = 1, 2$) if recipients in their village report being informally taxed during disbursement rounds one and/or two for the purposes of redistribution to non-recipients. As in Table 2.5, Columns 2-5 are estimated by weighted least squares where the weights for treatment households equal one and the weights for control households are given by the normalized estimated odds of treatment $\widehat{\omega} = \widehat{P}/(1 - \widehat{P})$, where the normalization is over the entire sample for the given time horizon. Using the Crump et al. (2009) procedure, we trim 41 households in the upper tail of the estimated propensity scores, $\widehat{P}$, and do the same in the OLS regressions for comparability. Column 3 controls linearly for the propensity score and column 5 for a fifth-order polynomial in the propensity score allowing it to vary by treatment and control. Column 4 controls for all covariates $\mathbf{X}_h$ used to estimate the propensity score. The coefficients on these additional terms in columns 3-5 are suppressed for presentational purposes. All columns include province fixed effects. Standard errors are clustered by village and computed over the entire two-step process using a block bootstrap with 1000 repetitions.

2.8.5 Data Construction

Merging Across Survey Waves. The main panel dataset is constructed by merging household identifiers across the three waves of *Susenas* enumerated in early 2005, 2006, and 2007. We first construct a balanced two-year panel between 2005 and 2006 by matching along (i) province-district-subdistrict-village-sampling ID-household ID and (ii) household head names in the 2005 and 2006 *Susenas* panels. While a traditional merge along strict geographic identifiers provides a balanced panel of $9,797$ households, significant discrepancies in household characteristics (including first names of household members) across waves indicate that survey administrators did not ensure the time-consistency of household presence in the physical location of prior enumeration. A name-matching algorithm provided by Robert Sparrow generated an initial balanced panel of almost 8500 households, and through further manual inspection, we added an additional 548 households. We then merged this balanced two-year panel of 9,048 households with the subset of all households enumerated in early 2007 also reporting enumeration in 2005 (in a specific survey question verified by enumerators). This process generated a balanced three-year panel comprising 7,014 households. A number of households are lost between 2006 and 2007 as a result of a change in the sampling rather than direct attrition by households.

Attrition. The data structure pose a nonstandard attrition problem. Although attritors appear much more similar to non-recipients than recipients (see Table 2.1), we do not know which attritors between 2005 and 2006 actually received the UCT. We observe recipient status among the 2,034 attritors between 2006 and 2007, and reassuringly the ratio of recipients to non-recipients remains essentially unchanged across years. Although inter-survey attrition is potentially a non-negligible problem, we ignore its consequences in the main results presented below. Nevertheless, all results are robust to reweighting the sample so as to account for the probability of attrition as a function of all observable characteristics used to predict treatment. The lack of contrast between attrition probability-weighted and unweighted estimates provides evidence against attrition being a source of misspecification that might confound identification of the causal average partial effect of interest (see Solon et al., 2013).[42]

Variable Construction. We use the *Susenas* datasets to construct all of the expenditure outcomes, baseline controls (see Table 2.2), health shocks and participation in other social programs (see Table 2.10), and information on UCT benefits.

The following variables were constructed based on data recorded in the Village Potential or *Podes* administrative census from April 2005: bank presence in the subdistrict, log travel distance to the (sub)district capital, log travel distance to the post office, village road paved, village accessibility by land, the incidence of natural disasters by type, and the ranking of the President's political party in the 2004 local village election. We are unable to match the data in *Podes* to *Susenas* villages for a small number of households in Papua. However, the main baseline results in Table 2.5 are robust to dropping these households.

Data on rainfall shocks were obtained from NOAA/GPCP sources. The total amount of rainfall during the given growing/harvest season where (i) seasons are 12 month intervals beginning with the first month of the province-specific wet season in a given year, and (ii) rainfall at the village level is based on rainfall levels recorded interpolated down to 0.5 degree (latitude/longitude) pixels between rainfall stations.

We obtain data on the regional CPI and the district-specific poverty lines (see Table 2.10) from the Central Bureau of Statistics. The former measures were mapped to villages based on the great circle distance from the centroid of the village's district to the centroid of the city in which the CPI measure was taken.

The (log) distance to Jakarta was computed as the great circle distance from the centroid of each district to the centroid of Jakarta

[42]A similar argument can be made for why we prefer not use the sampling weights provided in *Susenas*. Although we do not use those weights in the paper, doing so leaves key results are unchanged, and the lack of contrast between weighted and unweighted estimates provides evidence against the sort of misspecification that weighting might help correct in our context.

Chapter 2, in part, has been submitted for publication of the material, Bazzi, Samuel; Suryahadi, Asep; Sumarto, Sudarno. The dissertation author was the primary investigator and author of this paper.

Chapter 3

Blunt Instruments: Avoiding Common Pitfalls in Identifying the Causes of Economic Growth

Abstract

Concern has intensified in recent years that many instrumental variables used in widely-cited growth regressions may be invalid, weak, or both. Attempts to remedy this general problem remain inadequate. We show how a range of published studies can offer more evidence that their results are not spurious. Key steps include: grounding growth regressions in more generalized theoretical models, deployment of new methods for estimating sensitivity to violations of exclusion restrictions, opening the "black box" of GMM with supportive evidence of instrument strength, and utilization of weak-instrument robust tests and estimators.

3.1 Introduction

One of the great projects of economic research is to establish the causes of growth. Separating causes from correlates, however, is difficult. Many researchers have recently addressed this difficulty by deploying instrumental variables in cross-country datasets. This can help to identify causes of growth if the instruments do not materially affect growth through channels other than the variable of interest (the instruments are "valid") and if the instruments correlate well with the variable of interest (the instruments are "strong"). Unfortunately, for reasons not always transparent in published studies, these instruments can be invalid, weak, or both.

In this paper, we examine problems of instrument validity and strength in several growth papers recently published in general-interest and top field journals—not to single out those papers, but to concretely illustrate a general phenomenon that goes well beyond them. First, we discuss how an instrument that is plausibly valid when used in a single setting can be shown invalid by its use in additional settings. Second, we offer evidence that unacknowledged weak instruments may generate spurious findings in important applications, especially those using the popular Generalized Method of Moments (GMM) dynamic panel estimators. This evidence consists of simulation exercises and simple diagnostics on the data underlying published studies.

Our contribution is to show that these problems can have important consequences in real published work, and to suggest remedies. We advocate four ways that growth researchers can surmount these difficulties: by basing instrumental variable regressions on theory sufficiently general to comprise other published results with the same instrument, by using the latest methods to probe sensitivity to violations of the exclusion restriction, by opening the "black box" of GMM with complementary methods to assess instrument strength, and by deploying weak instrument robust testing procedures and estimators. We discuss each in detail below.

3.2 Instrumentation and its discontents

The wave of international growth empirics begun by Baumol (1986) and advanced by Barro (1991) inspired early skepticism even from its own contributors:

> "Using these regressions to decide how to foster growth is ... most likely a hopeless task. Simultaneity, multicollinearity, and limited degrees of freedom are important practical problems for anyone trying to draw inferences from international data. Policymakers who want to promote growth would not go far wrong ignoring most of the vast literature reporting growth regressions." (Mankiw, 1995).

Researchers thereafter began to address many of these problems. They became more assiduous in checking the robustness of results to the choice of regression specification (Fernández et al., 2001; Sala–i–Martin, 1997; Sala–i–Martin et al., 2004). They explored concerns about parameter heterogeneity, measurement error, and influential observations (Temple, 1999; Hauk and Wacziarg, 2009). They expanded their samples as the succession of years and improvements in information technology have brought a flood of new data (Bosworth and Collins, 2003; Easterly et al., 2004).

Beyond this, researchers have taken greater care in identifying the causal portion of the relationships they observe across countries. Architects of growth regressions published in top journals have used cross-country instrumental variables for governance quality,[1] trade,[2] and foreign aid,[3] among several other growth determinants. Advances in econometrics have assisted this search for better identification—especially the advent of sophisticated dynamic panel Generalized Method of Moments (GMM) estimators, which entered the growth literature with Caselli et al. (1996).

[1]These include cross-country instrumental variables based on exogenous deaths of national leaders while in office (Jones and Olken, 2005), colonial-era settler mortality (Acemoğlu et al., 2001), a Soviet-era survey of ethnolinguistic fractionalization (Mauro, 1995), distance from the equator (Hall and Jones, 1999), and Pacific-basin wind patterns (Feyrer and Sacerdote, 2004).

[2]These include cross-country instruments based on geography (Frankel and Romer, 1999; Frankel and Rose, 2002).

[3]These include cross-country instruments based on political ties, economic policies, and country size (Burnside and Dollar, 2000; Angeles and Neanidis, 2009). Boone (1996) also uses instruments based on political ties and country size in related work examining the impact of aid on a range of macroeconomic and development outcomes.

But in parallel with these welcome efforts, the economics literature in general has showed increasing concern with the strength and validity of instrumental variables in practice (surveyed by Murray, 2006). Close investigations have suggested that many cross-country instruments may be weak, invalid, or both, in widely-cited studies on the growth effects of governance or trade (e.g. Rodríguez and Rodrik, 2001; Brock and Durlauf, 2001; Dollar and Kraay, 2003; Glaeser et al., 2004; Albouy, forthcoming; Kraay, 2008). Notwithstanding the popularity of instrumental variables in recent growth empirics, Durlauf et al. (2005) conclude that "the belief that it is easy to identify valid instrumental variables in the growth context is deeply mistaken. We regard many applications of instrumental variable procedures in the empirical growth literature to be undermined by the failure to address properly the question of whether these instruments are valid". Acemoğlu (2010) decries the widespread use of "instruments without theory," and Hauk and Wacziarg (2009) see "unjustified claims of causality" as a prominent feature of growth empirics.

This paper extends a growing body of research aimed at identifying econometric best practice in growth empirics. First, we provide concrete evidence on ways in which published studies can collectively invalidate the instruments used in each study separately. Second, building on Bun and Windmeijer (2010) and Hauk and Wacziarg (2009), we indicate and suggest remedies for different sources of bias in the most popular estimator deployed in panel data growth econometrics, the system GMM estimator of Blundell and Bond (1998). Through a Monte Carlo simulation and simple diagnostic tests, we demonstrate the ways in which plausibly valid instruments can mask important weak instrument biases. We conclude with a discussion of solutions that applied researchers can deploy when faced with these identification challenges.

3.3 When strong instruments are invalid

To pass a rigorous peer review, each growth study employing an instrumental variable offers theoretical and empirical reasons to believe that the instrument is not substantially correlated with the regression's error term.

It is well known that this is difficult to establish. There can be a multiplicity of theoretical arguments for and against any given exclusion restruction, the true error term is unobserved in all applied settings, and empirical tests of overidentifiying restrictions, which often have low power, hinge on the untestable assumption that at least one instrument is valid—among other reasons.

What is not as well known is that *collectively* the literature establishes the invalidity of some instruments that growth econometricians now use widely, calling into question broad classes of their findings. Suppose that growth is determined by

$$g = \beta_0 + \sum_{j=1}^{k} \beta_j x_j + \varepsilon, \tag{3.1}$$

where g is growth, the x_j are a set of k potentially endogenous determinants of growth, the β are parameters to be estimated and ε is an error term with mean zero. Suppose we have an instrumental variable z such that $E[z\varepsilon] = 0$ but $Cov(z, x_j) \neq 0 \; \forall j$. We now try to estimate k separate regressions

$$g = \beta_j^0 + \beta_j x_j + \varepsilon_j, \quad j = 1, ..., k \tag{3.2}$$

in each case instrumenting for x_j with z where $\varepsilon_j \equiv \sum_{\ell \neq j} \beta_\ell x_\ell + \varepsilon$. But unless for every j it is the case that $\beta_\ell = 0$ (or more implausibly $x_j \approx x_\ell$) for all $\ell \neq j$, we have $Cov(z, \varepsilon_j) = \sum_{\ell \neq j} \beta_\ell Cov(z, x_\ell) \neq 0 \; \forall j$, and the instrument z is invalid in *every* regression (3.2). In other words, if existing research has shown that z is a strong instrument for a variable x_ℓ not included in a regression of the form (3.2) and $\beta_\ell \neq 0$, then z need not be a valid instrument for x_j. Any estimate $\widehat{\beta}_j$ will be biased to an unknown degree in an unknown direction, throwing into question the credibility of all results from the regressions (3.2). As Durlauf et al. (2005, p. 635) point out, "Since growth theories are mutually compatible, the validity of an instrument requires a positive argument that it cannot be a direct growth determinant or correlated with an omitted growth determinant."

And the story gets worse. We might think that including some of the omitted $x_{\ell \neq j}$ in the regression (3.2) could help, but that brings a new problem: For each $x_{\ell \neq j}$ included in (3.2), an additional instrument $\widetilde{z}$ is required—one that is valid ($E[\widetilde{z}\varepsilon] = 0$) *and* remains strong when used with the other (i.e., $Corr(z, x_j|\widetilde{z}) \neq 0$ and $Corr(\widetilde{z}, x_{\ell \neq j}|z) \neq 0$). This is a high bar.[4] Setting aside the difficulty of finding multiple valid instruments, Dollar and Kraay (2003) describe a case where each of two instruments appears strong in isolation but is so highly correlated with the other that both are weak when used together. We return to problems of instrument weakness in section 3.4.

3.3.1 Original sins

These systematic problems with instrument validity arise prominently in the widespread use of "legal origins" in growth regressions, a practice that has become the subject of frequent grumbling at conference coffee breaks. A flotilla of recent cross-country growth regressions has employed an indicator of the origin of a country's legal system (British, French, Scandinavian, and so on) as an instrument in a variety of regression specifications—each one of which suggests that the instrument is invalid in all of the other specifications. Many have passed the rigors of peer review at general-interest journals and top field journals.

Friedman et al. (2000) use legal origin as an instrument for five separate measures of "the quality of economic institutions" (corruption, tax rates, overregulation, etc.) in regressions with the size of the unofficial economy as the dependent variable—which could directly affect growth. Djankov et al. (2003) use legal origin as an instrument for "the degree of formalism of the legal procedure", which they argue causes a decline in the quality of the legal system (its honesty, impartiality, ability to enforce contracts, and so on) that could be a major determinant of growth. Lundberg and Squire (2003) use legal origin as an instrument for inflation, the inequality of land ownership, and several other variables that they argue directly affect growth. If any two of these studies are correct, growth is determined by a form of equation (3.1) that renders instrumentation in the IV regressions (3.2) invalid.

[4]In fact, nonzero partial correlation is not enough. If the instruments z and $\widetilde{z}$ are weak, then even a small degree of endogeneity in the instruments could lead IV estimates to be more biased than OLS (Hahn and Hausman, 2005).

It does not stop there. Alfaro et al. (2004) use legal origin as an instrument for private sector credit, bank credit, and stock market capitalization, which they argue condition the effect of Foreign Direct Investment on growth. Levine et al. (2000) similarly use legal origin to instrument for three separate proxies for financial intermediation, all of which they argue cause economic growth. Glaeser et al. (2004) use legal origin as an instrument for "executive constraints" and average years of schooling in the population, with the level of income per capita as the dependent variable. Beck et al. (2005) use legal origin as an instrument for "the relative size of the small and medium enterprise sector," which could be associated with growth. There are other examples.

If two or more of the above endogenous variables sufficiently affect growth, then instrumentation can be valid in at most one of these studies, and at worst none.

3.3.2 Size matters—through various channels

We turn to another instrument in widespread use, and dwell on it at greater length because its problems are less broadly recognized. Several recent cross-country studies published in general-interest journals and top field journals rest their identification strategies on the correlation of population size with some endogenous variable. In each case, the authors give plausible reasons why population size is not only a strong instrument but uncorrelated with their regressions' error terms: growth regressions do not typically find population scale effects (Rose, 2006; Easterly, 2009).

When viewed collectively, however, these studies exhibit a problem that undermines their careful arguments in support of instrument validity: Given that none of these studies include the other studies' endogenous variables as regressors, if population size is a strong and valid instrument in even one of these studies, then it is invalid in *all* of the others.[5] In other words, the conjecture in Deaton (2010) that measures of country size can affect growth through multiple channels

[5]Even if these studies included one or more of the endogenous variables in other studies, the authors would face precisely the problem discussed earlier, in the paragraph prior to Subsection 3.3.1.

has empirical support.

This pattern emerges in several recent and prominently published regressions. Some investigators use population size (among other geographic characteristics) as an instrument for trade as a determinant of the level of income per capita (Frankel and Romer, 1999; Frankel and Rose, 2002) or its growth (Spolaore and Wacziarg, 2005). Others regress growth not on the level of trade but on an indicator of the mix of goods exported, instrumented by population size (Hausmann et al., 2007), without controlling for the level of trade. Still others use population size as an instrument to identify the effect of foreign aid on democracy (Djankov et al., 2008), which many studies find to correlate with growth in some fashion.[6] Another approach uses country size—measured by area and level of GDP, but strongly correlated with population—to instrument for receipts of foreign direct investment (FDI) as a determinant of growth (Borensztein et al., 1998).

The exclusion restriction necessary for population size to be a valid instrument for each of these endogenous variables is violated to a greater degree, to the extent that the causal pathway identified in any of the other studies is correct. Regardless of any theoretical and empirical case for instrument validity made by each paper in the group, population size can only be a strictly valid instrument in one of them at best, and none of them at worst. The degree to which each estimate is thereby biased could be small or large, but should not be ignored.

3.3.3 Strength in numbers, but not validity

The problem extends further than this, however, in a way that is not generally recognized. Many studies resort to multiple instruments, responding to criticism by pointing out that allegations of invalidity or weakness only apply to some of the instruments. It is common to gloss over the problem that the most valid instruments in the basket could be the weakest, and that the strongest could be the least valid.

Building on the above discussion of the population size instrument, it is pos-

[6] For investigations of the effect of democracy on growth, see Barro (1996), Tavares and Wacziarg (2001), Giavazzi and Tabellini (2005), Rodrik and Wacziarg (2005), Persson and Tabellini (2006), Persson and Tabellini (2007), and Papaioannou and Siourounis (2008).

sible for a study whose identification strategy appears to rely on multiple instruments to rely in fact entirely on population size. Rajan and Subramanian (2008) execute cross-section regressions of growth on foreign aid receipts, with aid instrumented by a variable constructed (in an auxiliary or zero-stage regression) from aid-recipient population size, aid-donor population size, colonial relationships, and language traits (see Appendix subsection 3.8.1). Rajan and Subramanian write, "Our instrument ... contains information that is not just based on recipient size" (footnote 16).[7] But the instrument contains, in fact, almost no information beyond the size of the recipient's population. In Rajan and Subramanian's data, for the period 1970-2000, the in-sample correlation of log population and the constructed instrument is −0.93. In the periods 1980-2000 and 1990-2000, this correlation is −0.95. In effect, Rajan and Subramanian are instrumenting for aid with population alone, though they recognize the problem with using population size as an instrument.[8]

This problem deserves additional discussion, since it is common in applied work to rest identification on a group of instruments without making explicit which of them bears the burden of identification and therefore the key burden of validity. Frankel and Romer (1999) demonstrate that their gravity-based instrument—also constructed in an auxiliary regression—contains information beyond country size by treating log population and log area as exogenous and hence including them in both the first-stage and the second-stage.[9] Taking this minimalist approach, we

[7]They justify this claim (in their table 5, panel C) by using one measure of country size (population) as an excluded instrument in the construction of their generated instrument ($\overline{a}_r$) and, in a robustness check, showing that $\overline{a}_r$ retains strength when a different measure of country size (land area) is used as an additional excluded instrument in the first stage. But the only way to accurately assess whether or not $\overline{a}_r$ contains information beyond population size is to test whether or not it retains significance when population itself is included as a separate instrument as we do here.

[8]"While a measure of country size could in itself be a plausible instrument, the reason not to make it the preferred one is that there is uncertainty whether it can satisfy the exclusion restriction; that is, a number of reasons can be advanced as to why a recipient's size would have an independent effect on growth." (Rajan and Subramanian, 2008, footnote 16).

[9]However, upon more rigorous examination of the exclusion restrictions implicit in this instrument, Frankel and Rose (2002) conclude that among the six plausibly exogenous geographic determinants of trade flows used to construct their predicted trade instrument, log population is the only one that violates the implicit overidentifying restrictions used in constructing the instrument. See footnote 15 of Frankel and Rose (2002). This result supports our claims in this section about the non-excludability of size. Debate over other aspects of the Frankel and Romer

explore in Tables 3.1 and 3.2 the role of population as an instrument using the original data of Rajan and Subramanian.

For each specification, we test for underidentification and for weak instruments. To test for underidentification, we report p-values for a test of the null hypothesis that the structural equation is underidentified based on a Lagrange-Multiplier (LM) test using the rank-based rk statistic due to Kleibergen and Paap (2006). A rejection of the null indicates that the smallest canonical correlation between the endogenous variables and the instruments is nonzero. However, nonzero correlations are not sufficient for strong identification. We therefore also report first-stage F statistics—Wald statistics based on Cragg and Donald (1993) and the Kleibergen and Paap (2006) generalization to non-i.i.d. errors—and associated p-values for weak-instruments hypothesis tests.

Following the diagnostic approach developed in Stock and Yogo (2005) and implemented in Yogo (2004), we report p-values for the null hypotheses (i) that the bias in the point estimate(s) on the endogenous variable(s) is greater than 10% or 30% of the OLS bias, or (ii) that the the actual size of the t-test that the point estimate(s) on the endogenous variable(s) equal zero at the 5% significance level is greater than 10 or 25%.[10] While it has become common practice in the empirical growth literature to report first-stage F statistics, the inferential implications often go unstated. By reporting p-values, we offer a probabilistic lens into the weak-instruments problem.

Table 3.1 shows that essentially all instrumentation power in the primary Rajan and Subramanian specification comes from the population instrument. Column 1 exactly reproduces a representative cross-section regression (Rajan and Subramanian Table 4, column 2). Instrumentation is very strong, as indicated by the tests for underidentification and weak instruments. Column 2 of Table 3.1 includes

specification can be found in Rodríguez and Rodrik (2001) and Noguer and Siscart (2005).

[10] These p-values are based on comparing the appropriately scaled large-sample versions of the Cragg-Donald and Kleibergen-Paap statistics to the critical values in Stock and Yogo (2005). Critical values have not been tabulated for the Kleibergen-Paap rk statistic since the specific thresholds depend on the type of violation of the i.i.d. assumption, which differ across applications. We follow others in the literature and apply the critical values tabulated for the Cragg-Donald statistic to the Kleibergen-Paap results (see Baum et al., 2007). See 3.8.2 for further details.

log population in the second stage, and instrument strength collapses. We fail to reject the null hypothesis that the structural equation is underidentified. Applying the conditional likelihood ratio (CLR) test of Moreira (2003), which is robust to weak instruments, we obtain an uninformative confidence interval on aid/GDP comprising the entire real line. Column 3 discards Rajan and Subramanian's constructed instrument altogether and uses log population alone as an instrument for aid, giving results nearly identical to those in column 1. Column 4 re-estimates the constructed instrument without the population size terms, and instrument strength is abysmally low. Table 3.2 shows only the first-stage F statistics from the Rajan and Subramanian cross-section regressions for 1970- and 1980-2000 (results are similar for 1990-2000): first in exact replication of their results, then with population terms deleted from the construction of their instrument, then with the instrument constructed based only on population and its interactions.

In all cases, we cannot reject that the structural equation is underidentified and aid is weakly instrumented when information about population is absent from the constructed instrument, and strongly instrumented when (only) those variables containing information about population are present. Moreover, using the CLR method of inference robust to weak instruments, it is not possible to rule out extremely large *or* extremely small negative *or* positive effects of aid on growth.

The Rajan and Subramanian cross-section method is indistinguishable from instrumenting exclusively with aid-recipient population. The subsequent discussion of the validity of any other variable in the instrument matrix, then, is not informative about the causal relationship between aid and growth. What matters is the validity of the instrument that strongly identifies causation. Since that is only country size, the Rajan and Subramanian analysis shares the same problem faced by the other papers resting on the population instrument: All of the aforementioned papers that use the population instrument invalidate its usage in these studies, since the regressions there do not control for the level of trade, the mix of goods exported, FDI, or democracy. And the Rajan and Subramanian exercise does not resolve important questions about the validity of the population instrument in all of the other papers that use it because those papers do not control for aid receipts.

Beyond generated instruments

This problem is, in fact, much more general than the use of instruments generated from auxiliary regressions. As an example, we consider two other prominent studies in the aid and growth literature: Burnside and Dollar (2000) and its highly-cited antecedent in Boone (1996). Table 3.3 examines the interplay of instrument strength and validity in each of these studies, which employ country size alongside several other instruments in a pooled 2SLS specification.[11] Again we test for underidentification (Kleibergen-Paap LM test) and weak instruments (Cragg-Donald and Kleibergen-Paap Wald stats). We also show Hansen's J tests of the null hypothesis that—roughly speaking—the instruments are valid.

Unsurprisingly, instrumentation is strong in columns 1 and 3, which replicate the studies' baseline specifications including the size instruments. However, when relaxing the excludability of log population in column 2, instrument strength collapses as Boone's political instruments identifying prominent donor-recipient relationships are weakly correlated with aid/GDP. Meanwhile, relaxing the excludability of the size instruments in Burnside and Dollar, we find in column 4 that the remaining policy instruments (see notes below the table) still explain some of the variation in aid/GDP as we reject the null of underidentification. Yet, the point estimate on aid/GDP has 10-30 percent of the OLS bias. Further relaxing the excludability of the policy instruments, instrument strength drops considerably in column 5 as the remaining political instruments identifying prominent donor-recipient relationships again prove to be weakly correlated with aid/GDP. We can thus conclude that the seminal aid and growth studies due to Boone and Burnside and Dollar suffer from the same identification challenges as their most recent successor, Rajan and Subramanian.

In Table 3.3, we also deploy several tests of overidentifying restrictions aimed at characterizing instrument (in)validity. In column 1, the p-value of 0.12

[11] Here, we extend and elaborate upon related points raised in Clemens et al. (2012), which provides more detail on these two seminal aid and growth studies.

for the Hansen (1982) test provides evidence against the null hypothesis that the full set of instruments in Boone is valid (or the model is correctly specified). We find similar evidence when comparing the Hansen (1982) test statistics with and without the size instruments.[12]

Moreover, by relaxing the excludability of population size in column 2, the p-value on the smaller set of (weak) political instruments triples. Treating the size instrument as strong and the political instruments as weak *a priori*, we fail to reject the validity of population size on the basis of a Hausman-type test (see notes to the table) for the validity of a strong instrument in the presence of weak instruments (Hahn et al., 2011), which delivers a p-value of 0.05.[13] Turning to columns 3-5 for Burnside and Dollar, the message is less clear. Yet, we do find relatively lower p-values for the difference-in-Hansen tests pertaining to the validity of the population size instruments. Taken together, these tests and the associated point and set estimates for aid/GDP provide additional evidence of the difficulties that arise when weak instruments are valid and strong instruments are invalid.[14]

This problem also extends beyond pooled cross-section models to dynamic panel regressions with numerous non-size-based instruments. As an example, we consider the 10 year panel regressions in Hausmann et al. (2007). The authors utilize two estimators: (i) a pooled 2SLS estimator with log population and log area as instruments, and (ii) the Blundell and Bond (1998) dynamic panel system GMM estimator with instrumental variables that include log population and log area as well as the standard set of lagged covariates employed in this popular estimation strategy (see Section 3.4 below for a detailed discussion of this estimator).

Table 3.4 demonstrates that the key dynamic panel result in Hausmann et al. (2007) hinges on the excludability of country size from the levels equation—despite plausibly valid moment conditions comprising lagged levels and differences of the endogenous variables. The statistically and economically significant effect of

[12]See Hayashi (2000, pp. 220, 232–4) for a discussion of these tests of overidentifying restrictions based on the difference-in-Hansen or C statistic.

[13]See the notes to Table 3.3. While informative as a heuristic test, the asymptotic properties of this test have been criticized by Guggenberger (2009) and the authors themselves.

[14]Of course, one must also recognize that these tests of overidentifying restrictions ultimately hinge on the untestable assumption that at least one of the instruments is valid.

time-varying export product diversity (initial EXPY) on economic growth is driven primarily by its covariation with slowly changing log population and time-invariant log area. Columns 1 and 2 replicate the results from columns 6 and 8 of Table 9 in Hausmann et al..[15] In column 3, we do not exploit moment conditions for log population and log area in the difference equation. Doing so leaves the (difference-in-)Hansen test statistics and inference largely unchanged. In column 4, we do not exploit moment conditions for the size instruments in the levels equation, and in column 5, we do not exploit these moment conditions in either equation. As we treat country size as non-excludable in increasingly more equations in the system, we are less likely to reject the null of valid identifying restrictions. Of course, there may be other circumstances in which population size and/or area could be excludable from one equation, neither, or both.

Not unlike the previous examples from the aid and growth literature, identification of the panel regressions in Hausmann et al. depends crucially on size-based instruments that, if they are valid in this setting, require causal pathways identified in other studies to be incorrect.[16] While Hausmann et al. mention this being a potential problem in their pooled 2SLS specification,[17] they do not consider how or why the system GMM estimator fails to solve the problem.

We can go beyond mere suspicion that residuals in some of these studies are correlated with the endogenous variables in the other studies. Table 3.5 shows this within the Rajan and Subramanian framework. Here we perform ten cross-section OLS regressions, each with a candidate growth determinant x_j ($j = 1, \ldots, 10$) on the left-hand side that has been omitted from the Rajan and Subramanian re-

[15] Despite utilizing their original Stata code and dataset, the system GMM replication in column 2 differs slightly albeit immaterially from the published results. See Appendix 3.8.6.

[16] As we show in Appendix 3.8.5, this same set of results does not hold in the longer, 5-year panel periodization in Hausmann et al. (2007). Given the higher frequency and additional periods in this specification, the variation in the system GMM instruments comprised of lagged levels and differences of endogenous covariates swamp the potentially non-excludable variation in country size. We cannot reject the null of valid overidentifying restrictions implied by the full instrument matrix or the size instruments alone.

[17] "The variables used as instruments [log population and log area] fail the overidentification test in columns (2) and (6) [pooled 2SLS], most likely because they are persistent series akin to country fixed effects in a panel. Reassuringly, columns (4) and (8) show that the GMM setup where lagged levels and differences are used as instruments passes both the overidentification test and exhibits no second order correlation" (Hausmann et al., 2007, footnote 9).

gressions, $x_j = \beta_j \ln population + \mathbf{Z}'\Theta_j + u$, where the $\mathbf{Z}$ are the second-stage regressors (including a constant) treated as exogenous by Rajan and Subramanian. The table reports the point estimate and standard error for β_j in each case, beginning with an estimate for Aid/GDP from the Rajan and Subramanian study. Log population has a statistically significant partial relationship with several variables that are plausible growth determinants, in addition to foreign aid. These include trade (Frankel and Romer, 1999), foreign direct investment (Borensztein et al., 1998), education expenditure (Bosworth and Collins, 2003), inequality (Forbes, 2000), government consumption (found to correlate with country size by Alesina and Wacziarg (1998), and acknowledged as a robust growth determinant by Sala–i–Martin et al. (2004)), alongside multiple others.[18]

3.4 When valid instruments are weak

So far we have discussed cases of (mostly) strong instruments whose invalidity is difficult to detect. We turn now to cases of plausibly valid instruments whose weakness is difficult to detect.

The advent of dynamic panel GMM has been a boon to growth empiricists. These estimators take advantage of moment conditions not exploited in earlier dynamic panel two-stage least squares (2SLS) estimators. Whereas the Anderson and Hsiao (1982) estimator, for example, only exploits a single lag of endogenous right-hand side variables as instruments, the GMM estimator of Arellano and Bond (1991) (hereinafter Arellano-Bond) exploits deeper lags beyond the first or second, zeroing out lagged values that would be treated as missing in Anderson and Hsiao's 2SLS framework. Arellano and Bond's estimator, sometimes referred to as "difference" GMM, thus provides additional overidentifying restrictions without sacrificing sample size. The related system estimator of Arellano and Bover (1995) and Blundell and Bond (1998) (hereinafter Blundell-Bond) imposes additional mo-

[18]A further complication arises when one considers relaxing the assumption of linearity in the endogenous variables of interest. Although one could construct nonlinear functions of valid instruments to meet the necessary rank conditions in specifications with endogenous quadratic or interaction terms, the larger instrument set often proves weak in practice. We explore this issue further in Appendix 3.8.4 using the Rajan and Subramanian framework.

ment conditions allowing one to exclude once- or twice-lagged differences from an additional estimating equation in levels.[19] Deeper lags are redundant given the Arellano-Bond moment conditions. Both estimators can accommodate additional instruments as well.

The general dynamic panel estimating equation is of the form

$$g_{i,t} = \beta \ln y_{i,t-1} + \mathbf{x}'_{i,t}\boldsymbol{\gamma} + \psi_i + \nu_{i,t} \tag{3.3}$$

where $y_{i,t-1}$ is GDP per capita in country i at time $t-1$ from the World Development Indicators or Penn World Table, $g_{i,t}$ is percentage growth $(\Delta \ln y_{i,t})$,[20] $\mathbf{x}_{i,t}$ is a vector of growth determinants, ψ_i is a country fixed effect, $\nu_{i,t}$ is an idiosyncratic shock, $t = 1, \ldots, T$, and $i = 1, \ldots, N$.

Arellano-Bond estimation transforms equation (3.3) into first-differences and exploits the moment conditions $E(\ln y_{i,t-j}\Delta\nu_{i,t}) = 0$ and $E(\mathbf{x}_{i,t-k}\Delta\nu_{i,t}) = 0$ for $t = 3, \ldots, T$, $j = 2, \ldots, t-1$ and $k = k', \ldots, t-1$. While researchers commonly instrument for the lagged dependent variable to address dynamic panel bias, most have a particular interest in some possibly singleton subset of growth determinants in $\mathbf{x}$. Here, the literature goes one of two ways. Some authors treat that specific subset as endogenous or predetermined, where $k' = 2$ and $k' = 1$, respectively. Others treat all elements of $\mathbf{x}$ as endogenous, in which case $k' = 2$. Another key choice concerns the number of moment conditions. Asymptotically, one would want to use the full set of lags, but as Roodman (2009b) and others show—and as we reaffirm below—such choices can have important finite-sample consequences.

Developed in response to the well-known weak instruments problem in difference GMM, the Blundell-Bond estimator augments the Arellano-Bond difference (DIF) equation with a levels (LEV) equation. Specifically, this popular esti-

[19] Caselli et al. (1996) and Levine et al. (2000) were respectively the first to employ the Arellano and Bond (1991) and Blundell and Bond (1998) estimators in the empirical growth literature.

[20] Strictly speaking, among the papers revisited here, Voitchovsky (2005) uses the dependent variable $\Delta \ln y_{i,t}$ while Levine et al. (2000), Hausmann et al. (2007), and Rajan and Subramanian (2008) instead use the (very closely related) period-average annual per capita growth rate. The two exceptions are Hauk and Wacziarg (2009) and DeJong and Ripoll (2006), whose regressand is the level $\ln y_{i,t}$, which is also amenable to a growth interpretation, given the inclusion of lagged log income on the right-hand-side.

mator exploits an additional set of moment conditions $E(\omega_{i,t}\Delta \ln y_{i,t-1}) = 0$ and $E(\omega_{i,t}\Delta \mathbf{x}_{i,t-1}) = 0$ for $t = 3, \ldots, T$ where $\omega_{i,t} = \psi_i + \nu_{i,t}$, and $\mathbf{x}_{i,t}$ is assumed to be endogenously determined. These moment conditions are valid under joint mean stationarity of the $\ln y_{i,t}$ and $\mathbf{x}_{i,t}$ processes but also under weaker albeit less plausible conditions (see Blundell et al., 2000). This provides exclusion restrictions (based on lagged differences) for the growth determinants in equation (3.3) in levels. In theory, these moment conditions offer a credible identification strategy for researchers aiming to test the canonical Solow growth model or to highlight a salient source of heterogeneity in growth rates across countries.[21]

Often, however, a crucial question goes unexplored in applications of this new econometric technology: How much of the variance in the endogenous variables is explained by the instruments? A standard test for weak instruments in dynamic panel GMM regressions does not currently exist, so measuring instrument strength empirically is nontrivial.[22] Until now, skeptical researchers have been mostly concerned with finite-sample biases stemming from weak instruments in the Arellano-Bond estimator and violations of the initial conditions assumption in the Blundell-Bond estimator.[23] What most have failed to address, however, is a potentially equally important problem, weak instruments in Blundell-Bond. Although generally thought to be more robust to weak instruments than difference GMM, recent work shows that this system estimator can also suffer from serious weak instrument biases (Hayakawa, 2009; Bun and Windmeijer, 2010). In practice, most applications of system GMM simply assume that instruments are strong. We argue that instrument strength is an empirical question that can and should be directly tested.

[21]Bond et al. (2001) characterize the appropriateness of the moment conditions in the context of estimating the Solow model, while Hauk and Wacziarg (2009) point out that, at least in theory, *exogenous* growth models do not necessarily prescribe the use of an instrumental variables framework.

[22]See Stock and Wright (2000) on why the weak-instrument diagnostics for linear instrumental variables regression do not carry over to the more general setting of GMM.

[23]Bobba and Coviello (2007), for example, demonstrate that the null result in Acemoğlu et al. (2005) is reversed upon augmenting the weakly instrumented difference estimator with the levels equation in the system estimator. By necessity, we discuss weak instruments in the DIF equation of the system estimator, but we explicitly leave the validity issue aside as it has been thoroughly addressed elsewhere (Roodman, 2009b; Hauk and Wacziarg, 2009).

Below we investigate instrument strength in a variety of applications of system GMM: first in simulated data, and then in several influential growth regressions recently published in top field and general-interest journals. We follow a simple approach to assessing instrument strength in dynamic panel GMM regressions advanced *analytically* by Bun and Windmeijer (2010) and Hayakawa (2009) and *heuristically* in various settings (Blundell and Bond, 2000; Dollar and Kraay, 2003; Roodman, 2009a; Newey and Windmeijer, 2009).

Specifically, we construct the GMM instrument matrix for the difference and levels equation of the system estimator and carry out the corresponding regressions using 2SLS.[24] This permits simple and transparent tests of instrument strength in a closely related setting. Blundell et al. (2000) demonstrate that the system estimator is a weighted average of the difference and levels equations with the weights on the levels equation moments increasing in the weakness of the difference equation instruments.[25] So, if instrumentation of contemporaneous differences by once, twice or multiply lagged levels is weak, and instrumentation of contemporaneous levels by lagged differences is weak, this casts great doubt on the ability of GMM estimators to yield strong identification as used in these settings. Formalizing this intuition, Bun and Windmeijer demonstrate the explicit connection between cross-sectional concentration parameters (from the familiar Stock and Yogo (2005) setup) and instrument strength in the panel 2SLS equations of the type that we

[24]For the DIF and LEV equations, this instrument matrix, originally due to Holtz-Eakin et al. (1988), takes the form (see Roodman, 2009a):

$$\underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \dots \\ \ln y_{i1} & 0 & 0 & 0 & 0 & 0 & \dots \\ 0 & \ln y_{i2} & \ln y_{i1} & 0 & 0 & 0 & \dots \\ 0 & 0 & 0 & \ln y_{i3} & \ln y_{i2} & \ln y_{i1} & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}}_{\text{DIF}}; \quad \underbrace{\begin{bmatrix} 0 & 0 & 0 & 0 & \dots \\ \Delta \ln y_{i2} & 0 & 0 & 0 & \dots \\ 0 & \Delta \ln y_{i3} & 0 & 0 & \dots \\ 0 & 0 & \Delta \ln y_{i4} & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}}_{\text{LEV}},$$

where, for presentational purposes, we restrict attention to the respective moment conditions $E(\ln y_{i,t-j}\Delta\nu_{i,t}) = 0$ and $E(\omega_{i,t}\Delta \ln y_{i,t-1}) = 0$ in a five-period panel.

[25]The authors make this simple yet powerful point in the panel AR(1) model without covariates. The system estimator delivers the autoregressive point estimate of $\hat{\alpha}_s = \tilde{\delta}\hat{\alpha}_d + (1-\tilde{\delta})\hat{\alpha}_l$ where $\hat{\alpha}_l$ is the point estimate from the LEV equation, $\hat{\alpha}_d$ is the point estimate from the DIF equation, and $\tilde{\delta} = \frac{\hat{\pi}_d'\mathbf{Z}_d'\mathbf{Z}_d\hat{\pi}_d}{\hat{\pi}_d'\mathbf{Z}_d'\mathbf{Z}_d\hat{\pi}_d + \hat{\pi}_l'\mathbf{Z}_l'\mathbf{Z}_l\hat{\pi}_l}$, where $\hat{\pi}_j$ are the equivalent first-stage estimates using the instruments $\mathbf{Z}_j$ for $j = l, d$ in the LEV and DIF equation respectively. Their familiar setup motivates our heuristic use of the 2SLS analogues.

estimate.

Extending the setup in Bun and Windmeijer (2010) to the common case of multiple endogenous variables, we examine whether the additional moment conditions used in system GMM are actually strong enough to compensate for the well-established weak instruments in difference GMM estimation of growth models. We appeal to the results of Blundell et al. (2000) and Stock and Yogo (2005) in justifying our extension of the AR(1) analytics to the case of multiple endogenous regressors. In particular, we do not examine the strength of identification in the individual first-stage regressions in isolation, but rather, we test whether the instruments jointly explain enough variation in the multiple endogenous regressors to identify unbiased causal effects in the structural equation (3.3). These multivariate versions of the tests described in Section 3.3.3 allow us to characterize under- and weak-identification in the Blundell-Bond estimator. The rank-based LM test for underidentification due to Kleibergen and Paap readily applies to the panel 2SLS context here. Bun and Windmeijer provide evidence that the weak-instruments testing methods derived in the cross-section "are also informative about absolute and relative 2SLS bias when exploiting the whole panel." Although these tests should be considered heuristic in the panel setting considered here, their use is certainly preferable to ignoring the problem.[26]

Despite the large number of instruments in several specifications considered below, the Stock and Yogo weak instruments tests offer a powerful diagnostic tool.[27] Their critical values are available for up to 100 instruments. Those critical values exhibit a slow rate of decay as instruments increase beyond 30 or 40, and numerical results suggest their procedure is consistent for any number of instruments (Stock and Yogo, 2005, p. 90). According to Stock and Yogo, "Viewed as a test, the procedure has good power, especially when the number of instruments is large."

[26]One reason for caution is that in panels of the type studied here, both cross-section heteroskedasticity and time-series heteroskedasticity are likely, which means that the conventional F statistic is problematic (see Bun and de Haan, 2010).

[27]Hall et al. (2008) develop a method for using the Stock and Yogo diagnostics to select optimal instruments in GMM regressions via examination of corresponding 2SLS regressions. Their results show that a Stock and Yogo pre-test can actually be more powerful than using weak-instrument robust inference procedures (see subsection 3.4.5) with the full set of possibly suboptimal instruments.

3.4.1 Monte Carlo results

Our first step is to show that the system GMM estimator can often have poor size and power properties, depending crucially on the extent of endogeneity and on the strength of instrumentation. This is a different focus than the simulations in Blundell et al. (2000) and Bun and Windmeijer (2010), though we follow their specification, which differs slightly from equation (3.3). We simulate:

$$\begin{aligned} y_{i,t} &= \beta y_{i,t-1} + \gamma d_{i,t} + \psi_i + \nu_{i,t} \\ d_{i,t} &= \zeta d_{i,t-1} + \theta_i + \phi_{i,t} \qquad\qquad t = 1, \ldots, 6;\ i = 1, \ldots, 100 \end{aligned} \tag{3.4}$$

where the initial conditions $y_{i,0} = \big(\psi_i + \big(\gamma\theta_i / \big(1-\zeta\big)\big)\big)/\big(1-\beta\big) + \nu_{i,0}$ and $d_{i,0} = \big(\theta_i\big)/\big(1-\zeta\big) + \phi_{i,0}$ are sufficient to impose mean stationarity, an assumption on which the consistency of the Blundell-Bond estimator is predicated. The errors are distributed as

$$\nu_{i,t}, \phi_{i,t} \sim N\Big(\Big(\begin{array}{c} 0 \\ 0 \end{array}\Big), \Big(\begin{array}{cc} \sigma^2 & \omega \\ \omega & \sigma^2 \end{array}\Big)\Big) \quad \text{and} \quad \psi_i, \theta_i \sim N\Big(\Big(\begin{array}{c} 0 \\ 0 \end{array}\Big), \Big(\begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array}\Big)\Big) \tag{3.5}$$

The correlation coefficient for the shocks is $\rho = \frac{\omega}{\sigma^2}$. All simulation results employ the Windmeijer (2005) two-step correction, cluster standard errors by groups i, include time dummies in all equations, treat $y_{i,t-1}$ and $d_{i,t}$ as endogenous, and include the full set of available lags in the difference equation instrument matrix.[28]

Figure 3.1 shows results from this simulation with $\gamma = 0.3$ and $\beta = 0.2$ based on 500 repetitions. The horizontal axis shows different assumed values of $\zeta \in \{0.1, 0.2, \ldots, 0.9\}$, indicating the persistence of d over time, and the vertical axis compares the estimated $\hat{\gamma}$ (solid black line) to the true γ (dotted red line). The

[28]Note that this setup is consistent with the growth formulation in equation (3.3) after simply "relabeling" equation (3.4): relabeling $y_{i,t}$ with $\ln y_{i,t}$, subtracting $\ln y_{i,t-1}$ from both sides, and relabeling β with $\tilde{\beta} = \beta - 1$. Our variance formulation is analogous to the factor loadings representation of endogeneity in Blundell et al. (2000).

dashed lines show the average 95% confidence interval on $\hat{\gamma}$ across all repetitions. The top part of the figure shows the results for the difference GMM estimator, the bottom part for the system GMM estimator. Each small panel of the figure shows a different combination of the extent of endogeneity $\omega \in \{-0.1, -0.5, -0.9\}$, and the shock variance $\sigma^2 \in \{0.1, 0.5, 1, 5, 10\}$.

The magnitude of σ^2, which implicitly captures the ratio of the variance in idiosyncratic shocks to the variance in country fixed effects, has fundamental effects on instrument strength. While the theoretical apparatus in Blundell and Bond (1998) presumes that $\sigma^2 = 1$, it is more likely that $\sigma^2 < 1$ for typical applications in the empirical growth literature including those we consider in the next section. That is, the time-invariant heterogeneity in income levels across countries is likely to swamp the within-country variation in idiosyncratic shocks. The theoretical channel from low σ^2 to weak instruments is borne out in Bun and Windmeijer (2010), among others.

The performance of the difference GMM estimators is poor. In no case does the estimate of $\hat{\gamma}$ both reject the hypothesis that $\gamma = 0$ and fail to reject the hypothesis that $\gamma = 0.3$. For the more negative values of ω, bias is so extensive that the true value of γ is often rejected. The downwardly biased difference GMM estimates are consistent with the Monte Carlo findings in Blundell and Bond (1998), although our results seem to imply biases even at quite low levels of persistence in d and y. The system GMM estimator performs better: The estimate of $\hat{\gamma}$ only rejects the true value when ζ is low, that is when d is not sufficiently persistent over time. It is able to reject the hypothesis $\gamma = 0$, but only for high levels of $\zeta > 0.6$, which is consistent with the original motivation for the system estimator in Blundell and Bond. Additionally, whereas the difference GMM estimator is unaffected by the magnitude of σ^2, the system GMM estimator performs poorly for low σ^2 when the degree of endogeneity is not extreme ($\omega > -0.9$).

Figure 3.2 suggests that these problems are related to problems of weak- or under-identification. Here, the vertical axis shows the p-value from the Kleibergen-Paap LM test of the null hypothesis that the 2SLS regressions for the difference and levels equations corresponding to Figure 3.1 are underidentified or rank-deficient.

In the upper part of Figure 3.2, current differences are instrumented by the same GMM instrument matrix of lagged levels used in the difference GMM estimates of Figure 3.1. In the lower part of Figure 3.2, current levels are instrumented by the same matrix of lagged differences used in the levels equation of the system GMM estimates in Figure 3.1. LM test p-values greater than 0.1 (or more conservatively 0.05) point to potentially severe under-identification.

A clear pattern emerges: When instrumentation is weak in the 2SLS equations of Figure 3.2, the performance of the corresponding difference and system GMM estimates of γ is poor in Figure 3.1. When instrumentation in the 2SLS levels equation is strong (e.g., when $\zeta > 0.6$ and $\sigma^2 > 1$), the estimates $\hat{\gamma}$ show excellent size and power properties. In Appendix 3.8.5, we repeat the same exercise with $\beta = 0.8$, so that y is more persistent over time, with essentially the same result.

Figures 1 and 2 are sobering: Under reasonable parameter assumptions, the system GMM estimator is capable of leading a researcher to spurious conclusions—that d does not cause growth when it does, or that d has a negative effect on growth when the true effect is positive. A major part of the problem appears to be that in many cases there is no good reason to believe that lagged levels of the regressors explain a large portion of the variance in current differences, or vice versa. In these simulation results this is transparent by construction. We now proceed to illustrate that this concern may be far from hypothetical and may apply to recently published growth regressions.

3.4.2 Financial intermediation: Abundant instruments versus strong instruments

Table 3.6 revisits the dynamic panel GMM results of Levine et al. (2000) using the original data.[29] Column 1 reproduces a representative regression of growth on "liquid liabilities" (their Table 5, column 1). Column 2 gives the results of the

[29]Levine et al. (2000) conduct similar regressions with three different endogenous measures of financial intermediation. See Appendix 3.8.5 for a discussion of similar results using the other two measures.

closest reproduction of this regression we could achieve using the authors' original dataset, and the results match relatively well.[30] Again we test for underidentification (Kleibergen-Paap LM test) and weak instruments (Cragg-Donald and Kleibergen-Paap Wald tests).

Column 3 carries out the same regression using simple pooled OLS. In columns 4 and 5, we purge the country fixed effects from the regression by first-differencing (FD) and within-transformation (FE). While weak instruments typically bias difference GMM estimates downward, Bun and Windmeijer (2010) demonstrate how system GMM estimates can be biased upward. This bias increases in the ratio of (i) the variance of the fixed effects to (ii) the variance of the idiosyncratic shocks. Recall that our simulation results in the preceding section similarly present the least biases for low values of this ratio, or high values of σ^2. In column 5, the estimated ratio of variances is approximately 5 (i.e., $\hat{\sigma}_\nu^2/\hat{\sigma}_\psi^2 \approx 0.2$ in equation (3.3)).[31] Column 6 regresses differenced growth on differenced regressors, instrumented by lagged regressor levels analogous to the difference GMM estimator. Both the Kleibergen-Paap LM test of underidentification and the Cragg-Donald and Kleibergen-Paap Wald-type statistics show that instrumentation is very weak, far too weak for instrumentation to remove a substantial portion of OLS bias.[32]

An additional problem lurks below the surface: The sample contains 74 countries, and 75 different instrumental variables are used in the system estimator.[33] The large number of instruments relative to the number of groups may

[30] Our replication uses the original `DPD96` Gauss program employed by the authors. The remaining specifications in the table use Stata software. The number of observations reported by Gauss differs from that reported by Stata (compare columns 2 and 3) for reasons discussed in Appendix 3.8.6 where we also provide details on our attempted replication of their results including a full elaboration of the point estimates suppressed in column 2 of Table Table 3.6.

[31] We estimate these variance terms using the Baltagi and Chang (1994) method, which typically exhibits superior finite sample performance in unbalanced panels such as those commonly used in the growth literature.

[32] It is worth noting that the extremely high p-values we obtain in a number of specifications (i.e., failure to reject the null of weak instruments) are not uncommon (see Yogo, 2004). Nor are they indicative of underpowered or biased tests as, for example, a p-value of one in a Hansen test of overidentifying restrictions may be in the presence of "too many instruments" (see Bowsher, 2002; Roodman, 2009b).

[33] In both the levels and difference equations, 35 lagged regressors are used as instrumental variables—one for each of the seven endogenous right-hand side variables in each of the five

actually result in a failure to expunge the endogenous components of the right-hand side variables, thereby biasing the coefficient estimates towards those from the OLS estimator (see Beck and Levine, 2004; Calderón et al., 2002; Roodman, 2009b). In the limiting case, a 2SLS regression that had one instrument for each observation would show strong instrumentation but would produce coefficients exactly equal to those produced by OLS, and would not address endogeneity bias at all. The problem is perhaps even more serious in panels in which the cross-sectional variation dominates the within variation, as is common in growth regressions. Until recently, the literature has offered little guidance on the appropriate number of instruments relative to the number of groups and time periods.

Roodman (2009b) discusses a practical method for addressing this problem of "too many instruments" in dynamic panel GMM estimation. He suggests first restricting the number of lagged levels used in the instrument matrix for the difference equation, but since Levine et al. (2000) restrict their original matrix to a single lag, we must try an alternative approach. By "collapsing" the instrument matrix, we can effectively combine the instruments into smaller sets while retaining the same information from the original 75 column instrument matrix. The "collapsed" matrix contains one instrument for each lag depth instead of one instrument for each period *and* lag depth as in the conventional dynamic panel GMM instrument matrix.[34] Roodman suggests that a liberal rule of thumb is to become concerned when the number of instruments is close to the number of groups, as in the present case. Column 7 shows the results with the instrument matrix collapsed. Again, we cannot reject the null of underidentification, and weak instruments imply sub-

periods—along with the 5 period dummies included in the main equation.

[34]Collapsing leads to the following changes in the general, full DIF and LEV instrument matrices in footnote 25,

$$\underbrace{\begin{bmatrix} 0 & 0 & 0 & \cdots \\ \ln y_{i1} & 0 & 0 & \cdots \\ \ln y_{i2} & \ln y_{i1} & 0 & \cdots \\ \ln y_{i3} & \ln y_{i2} & \ln y_{i1} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}}_{\text{DIF-Collapsed}}; \quad \underbrace{\begin{bmatrix} 0 & \cdots \\ \Delta \ln y_{i2} & \cdots \\ \Delta \ln y_{i3} & \cdots \\ \Delta \ln y_{i4} & \cdots \\ \vdots & \ddots \end{bmatrix}}_{\text{LEV-Collapsed}},$$

where the first column in DIF-Collapsed corresponds to the first lag collapsed across periods 3-5, the second column to the second lag collapsed across period 4-5, etc.

stantial bias of the 2SLS estimates relative to pure OLS. Weak identification is not an artifact of too many instruments. Instrumentation this weak—no matter how valid—is incapable of testing hypotheses about coefficients in the main regression.

To test for weak instruments in the system estimator, we must also examine the levels equation independently of but in the same manner as the difference equation. Columns 8 and 9 conduct this parallel exercise for the levels equation. Since the difference equation is so weakly instrumented, the burden of strong identification in the system estimator relies on the levels equation moments. In column 8, the level of growth is regressed on the level of the regressors in a two-stage least squares framework, instrumented by the same lagged differences as in the levels equation of the system GMM estimator. Once again, instrumentation is too weak to address any substantial portion of OLS bias, thereby casting doubt on the system GMM point estimate for liquid liabilities in column 2, which is remarkably close to that for the levels equation in column 8. Using a collapsed instrument matrix in column 9 leaves these primary conclusions unchanged.

3.4.3 Weak aid or weak instruments?

Table 3.7 repeats this analysis for an entirely different set of regressions. It revisits the dynamic panel results of Rajan and Subramanian (2008) using the original data. Columns 1 and 2 exactly replicate their main Arellano-Bond (Table 9, column 1) and Blundell-Bond (Table 10, column 1) results. Column 3 shows the simple pooled OLS result, which appears quite similar to the system estimate in the preceding column. Columns 4 and 5 purge country fixed effects from the regression in column 3 via first-differencing (FD) and within-transformation (FE) with results similar to those for the Arellano-Bond estimator in column 1. Given that the estimated ratio of the variance of the time-invariant individual effects to the variance of idiosyncratic shocks is around three in column 5, this evidence suggests that instrumentation in these dynamic panel GMM regressions may be too weak to improve upon OLS.

Following the approach above, in column 6 we estimate the difference component of the system estimator in a 2SLS regression with exactly the same sequential moment conditions. Using the Kleibergen-Paap LM test, we cannot reject the null of underidentification, suggesting that identification is too weak to conduct meaningful hypothesis tests based on the difference equation alone. Although the Kleibergen-Paap Wald statistic appears high and we can reject large relative OLS bias on the basis of Stock and Yogo diagnostics, the perceived strength turns out to be a statistical artifact of the large, unrestricted GMM instrument matrix. After collapsing the 120 column instrument matrix, the Kleibergen-Paap Wald statistic falls dramatically in column 7, and we cannot reject the null that the difference equation exhibits more than 30 percent of the OLS bias. Columns 8 and 9 repeat the same exercise for the levels equation of the system estimator. Column 8 demonstrates underidentification and weak instruments in the standard wide instrument matrix, and collapsing does little to help. These results suggest that the similarity between the biased OLS estimates in columns 3-5 and the dynamic panel GMM estimates in columns 1 and 2 is not a coincidence. Weak instruments in both the difference and levels equations render hypothesis tests on the system GMM point estimate for aid/GDP unreliable.

3.4.4 Beyond aid and credit

The findings above are not peculiar to the specifications used in these two studies. In this subsection, we further examine the weak instruments problem in other recently published empirical applications. The goal is to highlight features of the data and panel setup that give rise to different outcomes in terms of the strength of identification. Table 3.8 reports weak instruments diagnostics for the baseline system GMM specifications in four studies published within the last five years. As before, we emphasize not the particular findings of each study but rather the sources and quality of identification. For each study, we report results based on (i) the Kleibergen and Paap LM test for underidentification and (ii) the weak instrument tests based on the Kleibergen-Paap and Cragg-Donald Wald statistics for the 2SLS estimates of the difference (DIF) and levels (LEV) equations separately with the full and collapsed instrument matrices.[35]

The first panel of the table unpacks the Blundell-Bond estimates of the Solow growth equations in Table 13 of Hauk and Wacziarg (2009). They treat all augmented Solow regressors—physical capital, human capital, population and lagged income—as endogenous and instrument with the full set of available lags in the difference equation. In only one specification—the levels equation with a collapsed instrument matrix—do we fail to reject the null of underidentification. Yet, weak instruments still afflict the system GMM estimates, as we cannot reject the null hypothesis that the 2SLS estimates maintain a nontrivial portion of the OLS bias. Compared with results in the previous subsections, however, it seems that the instruments explain some of the variation in the four endogenous variables in the canonical Solow model estimated over a sufficiently long panel.

Next, we return to the Hausmann et al. (2007) results, examining their longer panel employing a five year periodization. The longer panel affords a richer degree of within-country variation.[36] We can reject the null of underidentification in both the difference and levels equations based on the full, unrestricted instrument matrix. However, we cannot rule out underidentification when collapsing the instrument matrix for the difference equation. Nor can we rule out that weak instruments leave much of the OLS bias in the four 2SLS specifications. This is concerning since plausibly invalid country size instruments account for a nontrivial amount of the instrument strength captured by the underidentification and weak instrument test statistics and, especially under failures of validity, the 2SLS bias can be worse than the OLS bias. Using the Hahn et al. (2011) test for instrument validity, we strongly reject the validity of the lagged difference in log population as an instrument in the 2SLS levels equation specifications.

The third paper we consider is due to Voitchovsky (2005) who analyzes the impacts of inequality on economic growth using a short unbalanced panel of

[35]Our replications of Hauk and Wacziarg (2009), Hausmann et al. (2007) and DeJong and Ripoll (2006) are exact, relying on the original data and code provided by the authors. Our replication of Voitchovsky (2005), for which original code is unavailable, yields slightly different results than the published version. See Appendix 3.8.6.

[36]The ratio of the variance in country fixed effects to the variance in idiosyncratic heterogeneity is approximately unity compared to the shorter panel with ten year periodization (see Table 3.4) where that ratio exceeds two.

OECD countries from 1970-1995. We report our diagnostics for the baseline specification in Column 4, Table 2 (Voitchovsky, 2005, p. 287), in which five growth determinants are treated as endogenous—lagged income per capita, contemporaneous investment, lagged schooling, lagged Gini coefficient, and lagged ratio of the 90/75th percentile of the income distribution.[37] The Kleibergen-Paap LM test suggests that the DIF equation is underidentified. Yet, the large Kleibergen-Paap Wald statistic allows us to reject the null that weak instruments bias the 2SLS point estimates. This seeming anomaly disappears when collapsing the instrument matrix and hence can be explained by the large number of instruments (24) relative to the number of countries (21). While we cannot reject that the original LEV equation is underidentified, collapsing the instrument matrix delivers an exactly identified IV equation in which the two instrumental variables pass the LM rank test.

Lastly, we examine the system GMM estimates from DeJong and Ripoll (2006), a study arguing that the relationship between trade openness and economic growth depends on initial income. We consider the authors' baseline estimates from the fourth column of Table 2 (p. 631). The regressions examine *eight* growth determinants: life expectancy, female schooling, male schooling, lagged income per capita, ad valorem tariffs (import duties as share of imports), tariffs$\times$initial income/capita, investment/GDP, and government spending/GDP. The first four growth determinants are treated as predetermined, and the latter four as endogenous. Regardless of the specification, we fail to reject that the structural equation is underidentified, casting doubt on the ability of the system GMM estimator to solve the weak instruments problem evident in these 2SLS regressions.[38]

Collectively, the simulation results and analysis of the six papers considered above sound a warning note about the credibility of unexamined growth empirics using difference and system GMM estimation. We have shown that with a weakly instrumented levels equation, system GMM estimates can exhibit biases of similar

[37] See Appendix 3.8.6 for a discussion of the instrumental variables used.

[38] Not unlike the Voitchovsky (2005) result, the large Kleibergen-Paap Wald statistic in the DIF equation disappears when using the collapsed instrument matrix and is incongruent with the more reliable underidentification test.

in magnitude to uncorrected OLS variants. However, unlike the initial conditions restrictions on which the system GMM estimator is predicated, weak instruments can be diagnosed and (partially) addressed in many settings.

3.4.5 Weak-instrument robust inference

In this final subsection, we go beyond documenting the pervasiveness of weak instruments in system GMM to characterize the implications for inference about parameters of interest. Using the Hausmann et al. (2007) and Hauk and Wacziarg (2009) studies as examples, we conduct weak-instrument robust inference on the 2SLS difference and levels equations.[39] With a single endogenous variable, the CLR approach of Moreira (2003), which we used in subsection 3.3.3, permits inference that is immune to the damaging effects of weak instruments. However, other methods are required for regressions with multiple endogenous variables.

Here, we utilize the Kleibergen (2002) testing procedure, which has better power properties than the conventional Anderson and Rubin (1949) test in the presence of many instruments—the norm in the dynamic panel context. We describe the testing procedure in detail in Appendix 3.8.3 Using the resulting K statistic, we can derive joint confidence sets for multiple endogenous variables. Although computationally intensive, the Kleibergen procedure is robust not only to (many) weak instruments but also to invalid instruments (Doko and Dufour, 2008). This is important given concerns about the validity of the moment conditions in the levels equation of system GMM.

Figure 3.3(a) plots two-dimensional weak-instrument robust confidence sets for subsets of the three endogenous variables in the Hausmann et al. (2007) five-yearly panel: log initial export diversity (EXPY), log human capital, and log initial GDP per capita. The 95 percent confidence ellipses in the graphs represent (approximately) the boundary of the maximal area level set over the third endogenous variable in the full three-dimensional confidence ellipsoids. In the top graph, we

[39] Given our illustrative purposes here, we use these two studies for reasons of computational practicality: their relatively small number of endogenous variables are more amenable to the test procedure we deploy here than the large number of endogenous variables in some of the other studies considered above.

cannot reject that both log initial export diversity and human capital have zero effect on economic growth in the 2SLS difference equation. Turning to the levels equation, however, we cannot reject the null hypothesis that log initial export diversity has a positive effect. The same general pattern for export diversity holds when examining the two-dimensional confidence ellipse with log initial GDP per capita in the bottom figure. On the basis of these figures, we conclude that the key system GMM point estimate in Hausmann et al. (2007) is robust to the weak instruments problem identified in Table 3.8. This is reassuring given that we could strongly reject underidentification of the levels equation using the Kleibergen-Paap LM test.

In Figure 3.3(b), we plot two-dimensional weak-instrument robust confidence sets for subsets of the four endogenous variables in Hauk and Wacziarg (2009): physical capital, human capital, population and lagged income. We cannot reject that log human capital and log physical capital have null effects on economic growth in both the 2SLS difference and levels equations. The bottom graph reaffirms the null result for human capital when taking a different two-dimensional representation of the four-dimensional confidence ellipsoid. While the levels and difference equations are not underidentified (see Table 3.8), weak instruments may have implications for inference in the augmented Solow model.

3.5 Lessons

We demonstrate that invalid and weak instruments continue to be commonly used in the growth literature. This suggests that the warnings of Durlauf et al. (2005) and others on this subject have gone unheard. Weak and/or invalid instruments do not assist researchers in conducting meaningful hypothesis tests about the causes of growth. Continued use of problematic instruments in the growth literature risks pushing all of its findings further towards irrelevance.

Many of the papers discussed here contain explicit policy implications based on their results. Without strong and valid identification of causal relationships, such exercises may or may not carry policy implications, and require further investigation. Nevertheless, these studies remain valuable contributions to the literature for other reasons—especially their innovations in method.

We certainly do not recommend that economists refrain from pursuing pressing research questions until perfect methods arrive. But we suggest a handful of guidelines for the next generation of growth empirics:

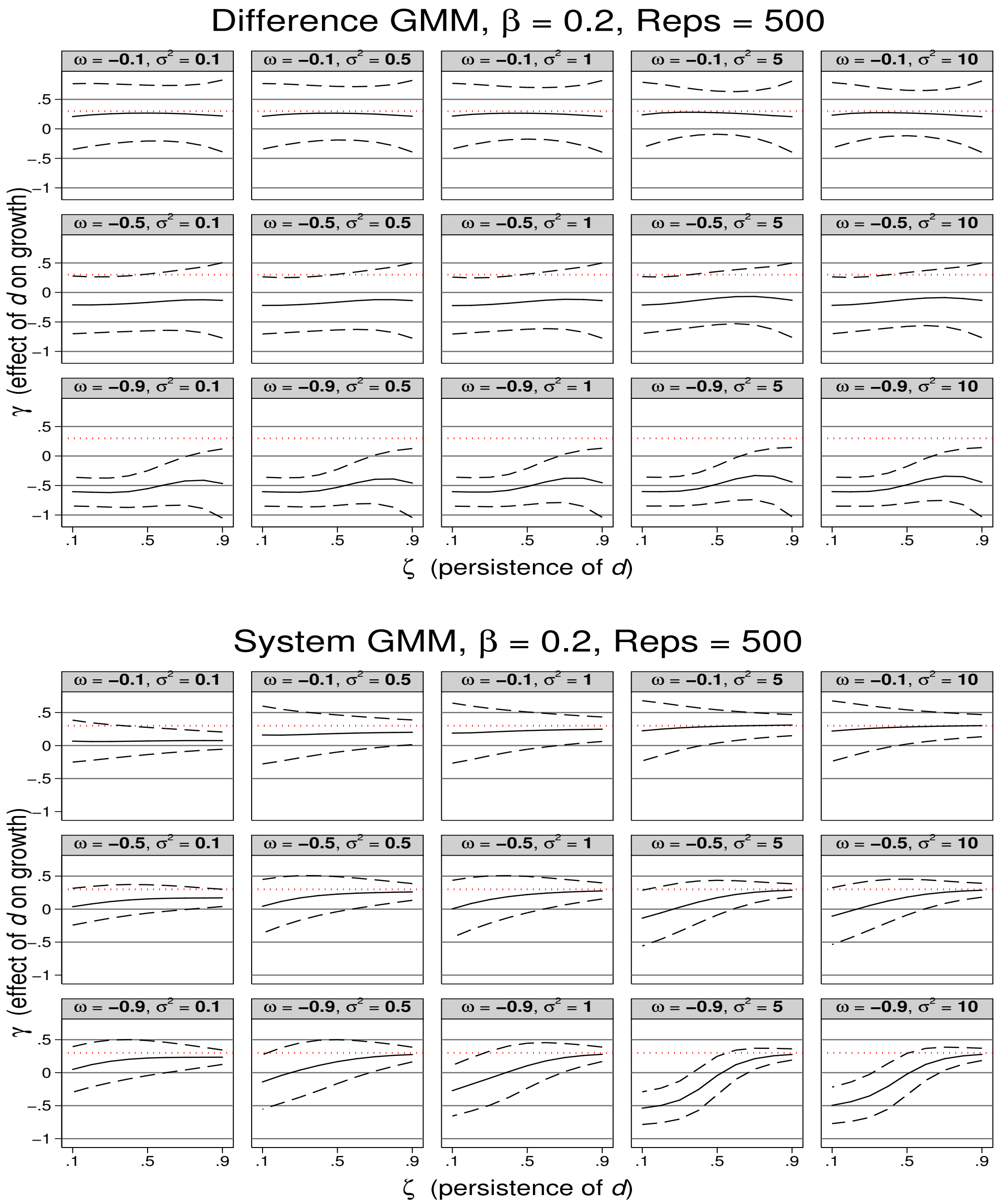
1. *Generalize the theoretical underpinnings of an instrument to account for other published results with the same instrument.* When an instrument has been used elsewhere in the literature, new users of that instrument bear the burden of showing that other important findings using that instrument do not invalidate its use in the new case. This can be done using a somewhat more generalized model that comprises causal pathways explored elsewhere with that instrument. Accounting for all plausible pathways through a "unified growth theory" is too high a standard, but accounting for prominent published pathways should be a minimum standard.

2. *Deploy the latest tools for probing validity.* Perfect instruments for growth determinants will remain elusive, but many underutilized tools exist to shine brighter light on the instruments we have. The Hahn et al. (2011) test used in this paper probes the validity of strong instruments in the presence of other weak ones. Imbens (2003) lays out a transparent method of assessing the sensitivity of a growth effect estimate to a given degree of correlation between instrument and error. Kraay (2008) and Conley et al. (2012) explore how to conduct second-stage inference accounting for prior uncertainty about the excludability of the instrument. Ashley (2009) shows how the discrepancy between OLS and IV estimates can be used to estimate the degree of bias under any given assumption about the degree to which the exclusion restrictions are violated.

3. *Open the black box of GMM.* It is no longer sufficient to assert that the mere use of system GMM adequately addresses the risk of weak instrumentation in dynamic panel models. As applied econometricians wait for an analog of the Stock and Yogo (2005) weak instrument diagnostics suitable for dy-

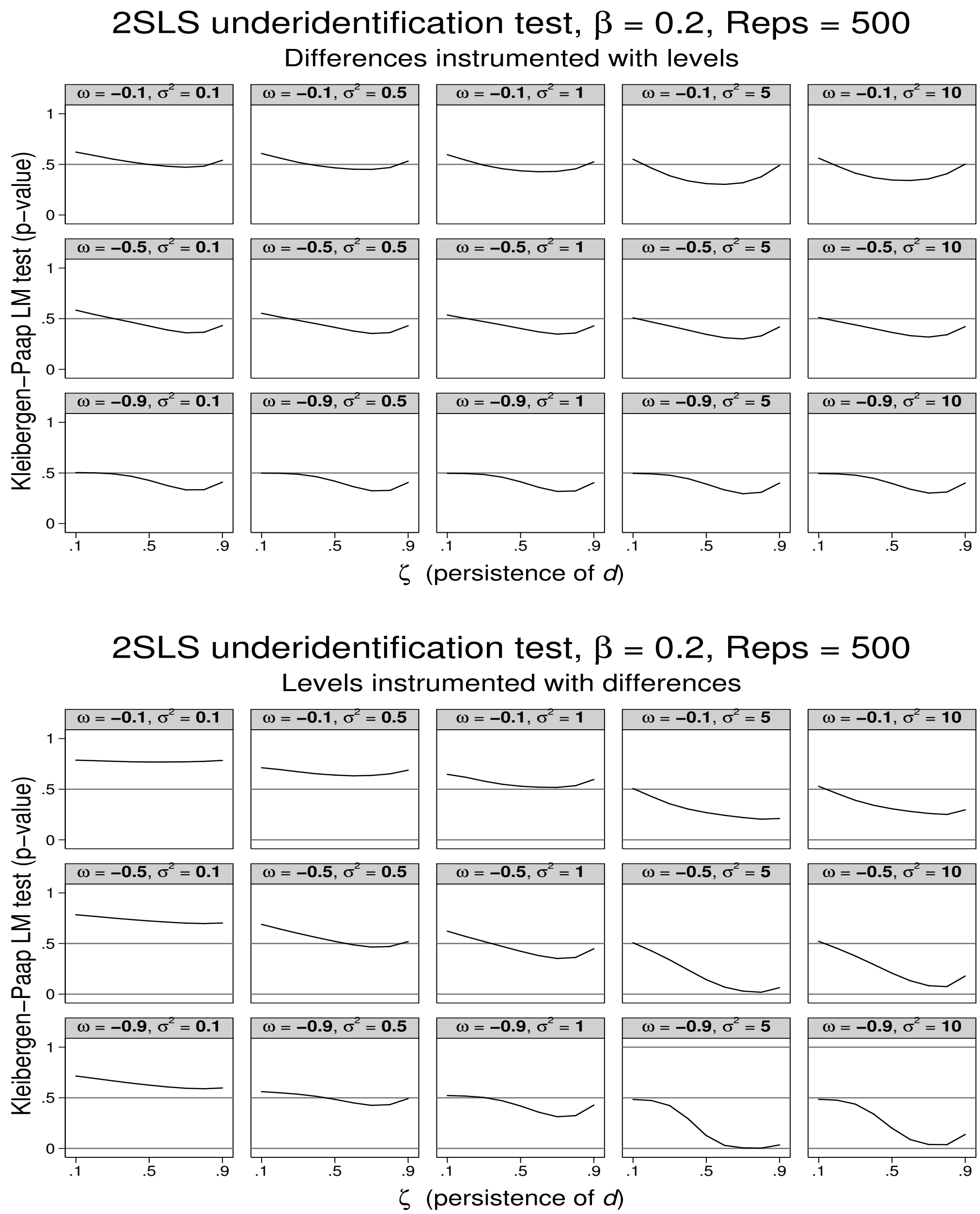
namic panel GMM estimation, its use must be complemented by supportive evidence that the instruments explain a sufficient degree of the variance of the endogenous regressors (and not simply because so many instruments are used). Papers exploring growth determinants should explore the strength of candidate instruments in analogous two-stage least squares regressions, should explore robustness to collapsing of the instrument matrix, should utilize optimal instrument selection procedures tailored to dynamic panel GMM (Okui, 2009), and should explore methods robust to weak instruments (Kleibergen, 2002; Kleibergen and Mavroeidis, 2009). Robust inference procedures now provide growth researchers with the means to go beyond merely identifying weak instruments to characterizing their implications for inference about key structural parameters.

3.6 Figures



Notes: The graphs show parameter estimates and 95% confidence intervals from simulations of the model in equation (3.4) based on 500 draws of a sample size of 600 with 100 cross-sectional units and 6 time periods, fixed $\beta = 0.2$, varying $\zeta \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$, varying degrees of endogeneity $\omega \in \{-0.1, -0.5, -0.9\}$, and alternative variances of the idiosyncratic shock, $\sigma^2 \in \{0.1, 0.5, 1, 5, 10\}$, where the variance of cross-sectional heterogeneity is fixed at 1. The dashed red line shows the true value of $\gamma = 0.3$ in the simulations.

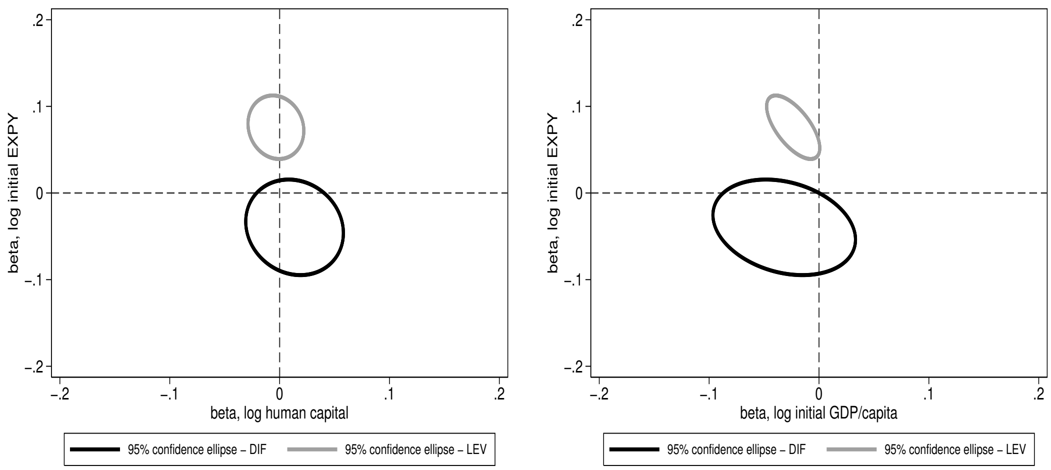
Figure 3.1: Power and size properties of GMM estimators in simulation results



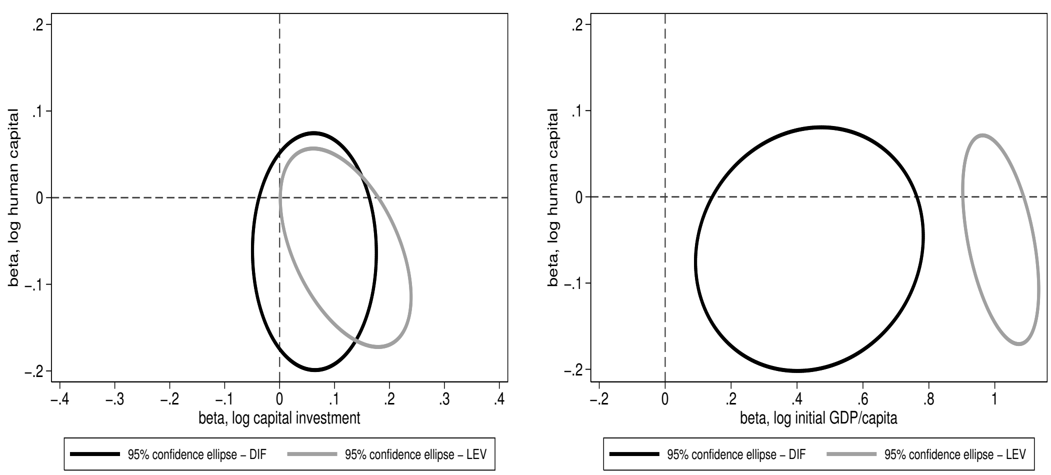
Notes: The graphs show p-values from a Kleibergen-Paap LM test for (the null of) underidentification in the levels and differences equations from simulations of the model in equation (3.4) as detailed in the notes to Figure 3.1. See the notes to Table 3.1 for details on the Kleibergen-Paap test.

Figure 3.2: Weak identification in simulation results, $\beta = 0.2$

(a) Hausmann et al. (2007)



(b) Hauk and Wacziarg (2009)



Notes: The graphs in (a) (quadrants I and II) are the 95% weak-instrument robust confidence ellipses for two of the three endogenous variables in the 2SLS analogues of the difference (DIF) and levels (LEV) equations in the system GMM estimates of the dynamic panel regressions in Hausmann et al. (2007). The confidence regions are obtained through a three-dimensional grid-search procedure over the domain -0.2 to 0.2 at increments of 0.01 for each of the three variables. The graphs in (b) (quadrants III and IV) are the 95% weak-instrument robust confidence ellipses for two of the three endogenous variables in the 2SLS analogues of the difference (DIF) and levels (LEV) equations in the system GMM estimates of the dynamic panel regressions in Hauk and Wacziarg (2009). The confidence regions are obtained through a four-dimensional grid-search procedure over a domain comprising the original point estimates and exceeding zero from above or below. The domain spans the $x-$ and $y-$axis for the variables shown in these graphs. The procedure is based on the approach developed in Kleibergen (2002) (see Appendix 3.8.3), which is shown to have higher power than the more familiar Anderson-Rubin test in the presence of many instruments. The ellipses are means-centered with a boundary constant of 4.

Figure 3.3: Weak-instrument robust confidence sets

3.7 Tables

Table 3.1: Rajan and Subramanian (2008) cross-section regressions, 1970–2000

	(1)	(2)	(3)	(4)
Point estimate: Aid/GDP	0.096	0.911	0.078	–15.944
	(0.070)	(4.083)	(0.066)	(633.474)
CLR confidence set: Aid/GDP*	[–0.027,0.291]	$(-\infty, \infty)$	[–0.039,0.252]	$(-\infty, \infty)$
Initial Log Population		1.604		
		(7.923)		
Initial Log GDP/capita	–1.409	1.061	–1.438	–25.491
	(0.435)	(12.782)	(0.403)	(953.378)
	Other parameter estimates omitted			
Excluded Instrument	$\overline{a}_r$	$\overline{a}_r$	ln(*population*)	$\overline{a}_r$ sans *population*
Observations	78	78	78	78
Kleibergen-Paap LM test (p-value)†	0.0004	0.772	0.0001	0.978
Cragg-Donald Wald stat‡	31.63	0.133	36.30	0.001
H_0: t-test size>10% (p-value)	< 0.001	0.982	< 0.001	0.999
H_0: t-test size>25% (p-value)	< 0.001	0.774	< 0.001	0.980
H_0: relative OLS bias>10% (p-value)	< 0.001	0.952	< 0.001	0.996
H_0: relative OLS bias>30% (p-value)	< 0.001	0.852	< 0.001	0.987
Kleibergen-Paap Wald stat‡	36.12	0.073	32.14	0.001
H_0: t-test size>10% (p-value)	< 0.001	0.987	< 0.001	0.999
H_0: t-test size>25% (p-value)	< 0.001	0.831	< 0.001	0.984
H_0: relative OLS bias>10% (p-value)	< 0.001	0.965	< 0.001	0.996
H_0: relative OLS bias>30% (p-value)	< 0.001	0.852	< 0.001	0.987

Notes: The dependent variable in all specifications is average annual growth in GDP per capita over the period. $\overline{a}_r$ is the generated instrument for foreign aid receipts/GDP (see Appendix 3.8.1). Heteroskedasticity-robust standard errors in parentheses. Column 1 exactly replicates the baseline result from Rajan and Subramanian (2008, Table 4, Column 2) for the 1970–2000 cross-section. Column 2 includes log population in the 2nd stage. Column 3 replaces estimated aid/GDP $\overline{a}_r$ with log population as the sole excluded instrument. Column 4 removes donor and recipient population terms from the zero-th stage specification used to estimate the predicted aid/GDP instrument $\overline{a}_r$, retaining only the colonial ties indicators. All specifications include dummies for sub-Saharan Africa and East Asia. *The CLR confidence set corresponds to the weak-instrument robust confidence set obtained using the conditional likelihood ratio test in Moreira (2003). †The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (i.e., the rank condition fails). The test uses a procedure from Kleibergen and Paap (2006). ‡In this special case of a single endogenous regressor, the Cragg-Donald and Kleibergen-Paap Wald statistics reduce respectively to the standard non-robust and heteroskedasticity-robust first-stage F statistics. Below each, we report the p-values from tests of whether (i) the actual size of the t-test that $\beta_{aid} = 0$ at the 5% significance level is greater than 10 or 25%, and (ii) the bias of the IV estimates of β_{aid} reported in the table are greater than 10 or 30% of the OLS bias. In both cases, the critical values are obtained from Stock and Yogo (2005). Although critical values do not exist for the Kleibergen-Paap statistic, we follow the approach suggested in Baum et al. (2007) and apply the Stock and Yogo critical values initially tabulated for the Cragg-Donald statistic. The critical values for (ii) are (less conservatively) based on three instruments since one cannot calculate critical values in the (finite-sample)bias tests for the case of one endogenous variable and fewer than three instruments.

Table 3.2: Instrument strength in Rajan and Subramanian (2008) cross-section

Period	1970–2000 ($N = 78$)			1980–2000 ($N = 75$)		
"Zero-Stage" Specification:	Replication	Colonial vars. only	Population vars. only	Replication	Colonial vars. only	Population vars. only
	(1)	(2)	(3)	(4)	(5)	(6)
Point estimate: Aid/GDP	0.096	–15.944	0.078	–0.004	–0.308	–0.028
	(0.070)	(633.474)	(0.067)	(0.095)	(0.389)	(0.084)
CLR confidence set: Aid/GDP	[–0.027,0.292]	$(-\infty, \infty)$	[–0.039,0.254]	[–0.186,0.232]	$(-\infty, \infty)$	[–0.194,0.170]
Kleibergen-Paap LM test (p-value)	0.0004	0.978	0.0001	0.0002	0.282	0.0001
Cragg-Donald Wald stat	31.63	0.001	35.90	29.37	1.41	40.54
H_0: t-test size>10% (p-value)	< 0.001	0.999	< 0.001	0.001	0.888	< 0.001
H_0: t-test size>25% (p-value)	< 0.001	0.980	< 0.001	< 0.001	0.341	< 0.001
H_0: relative OLS bias>10% (p-value)	< 0.001	0.996	< 0.001	< 0.001	0.772	< 0.001
H_0: relative OLS bias>30% (p-value)	< 0.001	0.987	< 0.001	< 0.001	0.503	< 0.001
Kleibergen-Paap Wald stat	36.12	0.001	31.62	31.26	1.41	39.65
H_0: t-test size>10% (p-value)	< 0.001	0.999	< 0.001	0.001	0.888	< 0.001
H_0: t-test size>25% (p-value)	< 0.001	0.984	< 0.001	< 0.001	0.340	< 0.001
H_0: relative OLS bias>10% (p-value)	< 0.001	0.997	< 0.001	< 0.001	0.770	< 0.001
H_0: relative OLS bias>30% (p-value)	< 0.001	0.990	< 0.001	< 0.001	0.502	< 0.001

Notes: In all specifications, the instrumental variable is aid/GDP predicted from the zero-stage regression. The dependent variable in all specifications is average annual growth in GDP per capita over the period. Heteroskedasticity-robust standard errors in parentheses. Following the original paper, we retain the degrees-of-freedom adjustment to the Kleibergen-Paap F and LM statistics based on robust standard errors. For each of the three periods, the first column is based on exact replication of the baseline result in Rajan and Subramanian (2008, Table 4); the second column removes donor and recipient population terms from the zero-th stage specification used to estimate the predicted aid/GDP instrument $\overline{a}_r$, retaining only the colonial ties indicators; the third column retains only the population terms in the zero-th stage. All specifications include dummies for sub-Saharan Africa and East Asia. See the notes to Table 3.1 for more details on the CLR confidence set as well as the Kleibergen-Paap and Cragg-Donald tests. Results for the period 1990-2000 are similar and can be found in Appendix 3.8.5.

Table 3.3: Unpacking the sources of identification in aid-growth regressions

Study	Boone (1996)		Burnside and Dollar (2000)		
Population IVs in 2nd stage[∓]	—	Yes	—	Yes	Yes
Policy IVs in 2nd stage[±]	—	—	—	—	Yes
	(1)	(2)	(3)	(4)	(5)
Point estimate: Aid/GDP	0.235	–0.782	–0.119	–0.206	0.363
	(0.198)	(0.818)	(0.180)	(0.441)	(1.190)
CLR confidence set: Aid/GDP	[–0.125,0.562]	$(-\infty, 0.261] \cup [8.674, \infty)$	[–0.523,0.231]	[–1.314,0.609]	[–1.923,3.582]
Observations	132	132	275	275	275
Hansen test all instruments (p-value)	0.123	0.368	0.194	0.290	0.122
Hansen test excl. size instruments (p-value)	0.197	—	0.313	—	—
Difference-in-Hansen test (p-value)[†]	0.112	—	0.154	—	—
Hansen test excl. policy instruments (p-value)	—	—	0.078	0.169	—
Difference-in-Hansen test (p-value)[†]	—	—	0.799	0.456	—
Hansen test excl. size & policy instruments (p-value)	—	—	0.237	—	—
Difference-in-Hansen test (p-value)[†]	—	—	0.230	—	—
Kleibergen-Paap LM test (p-value)	0.004	0.201	< 0.0001	0.057	0.124
Cragg-Donald Wald stat	15.70	1.77	19.74	7.04	4.65
H_0: relative OLS bias>10% (p-value)	0.010	0.724	< 0.001	0.510	0.409
H_0: relative OLS bias>30% (p-value)	< 0.001	0.442	< 0.001	0.014	0.162
Kleibergen-Paap Wald stat	7.57	1.69	15.76	5.90	2.77
H_0: relative OLS bias>10% (p-value)	0.283	0.734	0.001	0.687	0.603
H_0: relative OLS bias>30% (p-value)	0.028	0.455	< 0.001	0.043	0.313

Notes: The dependent variable in all specifications is average annual growth in GDP per capita over the period. The regressions replicated and modified are Boone (1996, Table 4, column V, row 3) and Burnside and Dollar (2000, Table 4, column 3 2SLS). The other coefficients are suppressed, but details on the replication of the original studies can be found in Clemens et al. (2012), which reported an abbreviated version of figures in this table. [∓]The log population instrument is included in the second stage of the Boone regression, and the following instruments are included in the second stage of the Burnside and Dollar regression: log population, log population×policy, and (log population)^2×policy. [±]The following instruments are included in the second stage of the Burnside and Dollar regression: log initial income×policy, (log initial income)^2×policy, and lagged arms imports/total imports×policy. [†]The null hypothesis of the difference-in-Hansen test (or C statistic, see Hayashi, 2000) is that the given suspect instruments are valid. This test is not robust to weak instruments. Applying an alternative Hausman test that is robust to weak instruments albeit problematic for other reasons (see Hahn et al., 2011), we fail to reject the null hypothesis that log population is a valid instrument in column 1 (p-value of 0.053) for Boone. Applying the same test to column 3 for Burnside and Dollar, we fail to reject that the size, policy, or size and policy instruments combined are valid with p-values of 0.411, 0.622, and 0.202 respectively. Following the original papers, we retain the degrees-of-freedom adjustment to the Kleibergen-Paap F and LM statistics based on country-level clustered standard errors in Boone and robust standard errors in Burnside and Dollar. See the notes to Table 3.1 for more details on the CLR confidence set as well as the Kleibergen-Paap and Cragg-Donald tests.

Table 3.4: Unpacking the sources of identification in Hausmann et al. (2007)

Estimator	IV∓	GMM-SYS∓	GMM-SYS	GMM-SYS	GMM-SYS	GMM-SYS
GMM-SYS Moment Conditions?	No	Yes	Yes	Yes	Yes	Yes
Size Instruments?	Yes	Yes	Yes, lev. eq.	Yes, diff. eq.	No	Yes
Size Excluded?	Yes	Yes	Yes	Yes	Yes	No
	(1)	(2)	(3)	(4)	(5)	(6)
log initial GDP/capita	–0.038	–0.013	–0.015	0.003	0.011	0.011
	(4.425)	(1.567)	(1.687)	(0.233)	(0.984)	(1.157)
log initial EXPY	0.092	0.043	0.047	0.008	–0.017	–0.017
	(4.598)	(2.315)	(2.600)	(0.213)	(0.777)	(0.796)
log human capital	0.004	0.005	0.004	0.000	0.007	0.005
	(1.766)	(0.652)	(0.515)	(0.024)	(1.242)	(0.959)
log area						–0.004
						(3.315)
log population						0.007
						(3.267)
Observations	299	299	299	299	299	299
Number of Periods	3	3	3	3	3	3
Number of Countries	79	79	79	79	79	79
Number of Instruments	2	18	18	18	16	18
Hansen test (p-value)	0.001	0.093	0.090	0.103	0.192	0.186
Hansen test excl. size instruments (p-value)		0.146	0.154	0.165	—	—
Difference-in-Hansen test (p-value)†		0.125	0.108	0.120	—	—
Kleibergen-Paap LM test (p-value)	< 0.001	—	—	—	—	—
Cragg-Donald Wald stat	17.47	—	—	—	—	—
H_0: t-test size>25% (p-value)	< 0.001	—	—	—	—	—
Kleibergen-Paap Wald stat	15.20	—	—	—	—	—
H_0: t-test size>25% (p-value)	< 0.001	—	—	—	—	—

Notes: The dependent variable in all specifications is average annual growth in GDP per capita over the period. The size instruments include log population and log area. The internal instruments refer to the lagged levels and lagged differences of endogenous right-hand side variables in the respective difference and levels equations of the dynamic panel GMM system of equations. ∓Columns 1 and 2 are based on Table 9, Columns 6 and 8 of Hausmann et al. (2007). We use Stata code and data provided by one of the authors, Jason Hwang, but the estimates in column 2 slightly differ from those reported in the published version of their paper. Following the original paper, we report heteroskedasticity-robust standard errors in parentheses and retain associated degrees of freedom adjustments for the first-stage test statistics. †The null hypothesis of the Difference-in-Hansen test is that the size instruments are valid. See the notes to Table 3.1 for more details on the Kleibergen-Paap and Cragg-Donald tests, which apply in column 1 to the endogenous log initial EXPY.

Table 3.5: log Population and omitted growth determinants

Dependent Variable	log Population regressor		
	Coefficient	*Std. Error*	*Observations*
Aid/GDP	–1.925	(0.340)	78
Trade/GDP	–13.680	(2.497)	77
FDI/GDP	–0.537	(0.183)	77
Education Expenditure/GDP	–0.423	(0.179)	75
Gini coefficient	–2.452	(0.991)	62
Government Consumption/GDP	–1.399	(0.352)	78
Manufacturing Value Added/GDP	1.529	(0.398)	76
Military Personnel/Total Labor Force	–0.263	(0.123)	78
Private Capital Flows/GDP	–2.548	(1.057)	77
Public Debt Service/GNI	–0.396	(0.229)	73
Savings/GDP	3.245	(1.502)	78

Notes: Each of the rows in the table correspond to a regression of the given dependent variable X listed in column 1 on log population and the additional covariates $\mathbf{Z}$ other than aid/GDP in the baseline 1970–2000 cross-section specification of Rajan and Subramanian (2008, Table 4, Column 2): $x_i = \beta \ln population_i + \mathbf{Z}_i'\mathbf{\Theta} + u_i$. Only the point estimates and standard errors for log population are reported. The standard errors are robust to heteroskedasticity. The sample sizes change depending on the number of available observations for the given dependent variable, all of which come from the World Bank's *World Development Indicators 2007* (Aid/GDP, Trade/GDP, FDI/GDP, Education Expenditure/GDP, Gini Coefficient, Government Consumption/GDP, Manufacturing Value Added/GDP, Military Personnel/Total Labor Force, Private Capital Flows/GDP, Public Debt Service/GNI, and Savings/GDP).

Table 3.6: Weak instruments in Levine et al. (2000)

	(1)	(2)	(3)	(4)	(5)
Estimator	GMM-SYS∓	GMM-SYS∓	OLS	OLS-FD	OLS-FE
Collapsed IV matrix	No	No	—	—	—
Liquid liabilities	2.952	2.834	1.692	1.095	0.851
	(0.001)	(0.001)	(0.000)	(0.122)	(0.296)
Log initial GDP/capita	–0.742	–0.792	–0.400	–13.609	–7.478
	(0.001)	(0.001)	(0.025)	(0.000)	(0.000)
	Other parameter estimates omitted				
Observations	359	359	345	323	345
Number of countries	74	74	74	74	74
Number of instruments	75	75	—	—	—
IV: Lagged levels	Yes	Yes	—	—	—
IV: Lagged differences	Yes	Yes	—	—	—

	(6)	(7)	(8)	(9)
	Difference Equation		**Levels Equation**	
Estimator	2SLS	2SLS	2SLS	2SLS
Collapsed IV matrix	No	Yes	No	Yes
Liquid liabilities	–0.747	–15.403	2.830	2.285
	(0.705)	(0.702)	(0.002)	(0.321)
Log initial GDP/capita	–12.435	–12.335	0.339	1.839
	(0.000)	(0.355)	(0.619)	(0.423)
	Other parameter estimates omitted			
Observations	323	323	345	345
Number of countries	74	74	74	74
Number of instruments	40	12	40	12
IV: Lagged levels	Yes	Yes	No	No
IV: Lagged differences	No	No	Yes	Yes
Kleibergen-Paap LM test (p-value)	0.780	0.580	0.559	0.200
Cragg-Donald Wald stat	0.59	0.04	0.72	0.25
H_0: relative OLS bias>30% (p-value)	1.000	1.000	1.000	0.997
Kleibergen-Paap Wald stat	0.67	0.06	1.12	0.25
H_0: relative OLS bias>30% (p-value)	1.000	1.000	1.000	0.998

Notes: The dependent variable in all specifications is average annual growth in GDP per capita each period. ∓Column 1 reproduces the published version of Levine et al. (2000, Table 5, Column 1), and column 2 reports our best attempted replication using the `DPD96` program for Gauss, the publicly available dataset, and a Gauss program used to generate their results provided by Thorsten Beck. Further details on the difference in sample sizes across columns, our replication efforts, and the associated differences in the Gauss and Stata programs for dynamic panel GMM regressions can be found in Appendix 3.8.6. The following variables are included in the regressions but suppressed in the table here for presentational purposes: government size, openness to trade, inflation, average years of secondary schooling, black market premium, time period dummies and a constant. The first five of these variables are treated as endogenous. Following the original paper, we report p-values in parentheses. See the notes to Table 3.1 for more details on the Kleibergen-Paap and Cragg-Donald tests, which apply in columns 6-9 to the full set of endogenous right-hand-side variables.

Table 3.7: Weak instruments in Rajan and Subramanian (2008)

	(1)	(2)	(3)	(4)	(5)
Estimator	GMM-DIF$^{\mp}$	GMM-SYS$^{\mp}$	OLS	OLS-FD	OLS-FE
Collapsed IV matrix	No	No	—	—	—
Aid/GDP	–0.151	–0.054	–0.037	–0.236	–0.224
	(0.077)	(0.114)	(0.053)	(0.066)	(0.067)
Initial log GDP/capita	–8.347	–2.456	–1.514	–13.245	–7.960
	(1.543)	(1.057)	(0.517)	(1.839)	(1.307)
	Other parameter estimates omitted				
Observations	359	359	345	323	345
Number of countries	74	74	74	74	74
Number of instruments	75	75	—	—	—
IV: Lagged levels	Yes	Yes	—	—	—
IV: Lagged differences	Yes	Yes	—	—	—
Lags used	2nd-7th	2nd-7th	—	—	—

	(6)	(7)	(8)	(9)
	Difference Equation		**Levels Equation**	
Estimator	2SLS	2SLS	2SLS	2SLS
Collapsed IV matrix	No	Yes	No	Yes
Aid/GDPs	–0.220	–0.355	0.116	0.470
	(0.086)	(0.157)	(0.079)	(0.710)
Initial log GDP/capita	–11.060	–10.535	0.117	10.193
	(1.980)	(3.355)	(1.454)	(15.689)
	Other parameter estimates omitted			
Observations	167	167	239	239
Number of countries	68	68	72	72
Number of instruments	120	52	41	17
IV: Lagged levels	Yes	Yes	No	No
IV: Lagged differences	No	No	Yes	Yes
Lags used	2nd-7th	2nd-7th	2nd	2nd
Kleibergen-Paap LM test (p-value)	0.522	0.698	0.765	0.413
Cragg-Donald Wald stat	0.66	0.43	0.41	0.06
H_0: relative OLS bias>30% (p-value)	1.000	1.000	1.000	1.000
Kleibergen-Paap Wald stat	9.89	1.36	0.69	0.07
H_0: relative OLS bias>30% (p-value)	< 0.001	1.000	1.000	1.000

Notes: The dependent variable in all specifications is average annual growth in GDP per capita each period. $^{\mp}$Column 1 exactly replicates Rajan and Subramanian (2008, Table 9, Column 1), and column 2 exactly replicates Table 10, Column 1 in Rajan and Subramanian. The following variables are included in the regressions but suppressed in the table here for presentational purposes: life expectancy, institutional quality, log inflation, M2/GDP, budget balance/GDP, revolutions, ethnic fractionalization, geography, time period dummies, dummies for countries in Sub-Saharan Africa East Asia, and a constant. The first six of these variables are treated as endogenous. Following the original paper, we report heteroskedasticity-robust standard errors in parentheses and retain associated degrees of freedom adjustments for the first-stage test statistics. See the notes to Table 3.1 for more details on the Kleibergen-Paap and Cragg-Donald tests, which apply in columns 6-9 to the full set of endogenous right-hand-side variables.

Table 3.8: Weak- and under-identification in dynamic panel growth regressions

Equation	Sample Size	No. of Endog. Vars.	No. of Instr-uments	KP LM test† (p-value)	CD F Stat‡	rel. OLS bias>30% (p-value)	KP F Stat‡	rel. OLS bias>30% (p-value)
	Hauk and Wacziarg (2009): panel with 69 countries, 8 periods							
DIF	414	4	102	0.075	1.84	1.000	4.34	0.002
DIF-Collapsed	414	4	27	0.012	1.67	0.995	2.43	0.861
LEV	483	4	27	0.042	1.74	0.992	1.83	0.987
LEV-Collapsed	483	4	4	0.161	0.94	0.783	0.49	0.913
	Hausmann et al. (2007)†: panel with 79 countries, 8 periods							
DIF	525	3	56	0.042	2.23	0.992	2.78	0.848
DIF-Collapsed	525	3	14	0.244	1.70	0.930	0.86	0.999
LEV	604	3	20	0.001	2.81	0.596	3.42	0.288
LEV-Collapsed	604	3	5	0.015	1.82	0.508	2.10	0.431
	Voitchovsky (2005): panel with 21 countries, 5 periods							
DIF	61	5	24	0.641	0.56	1.000	103.8	< 0.001
DIF-Collapsed	61	5	6	0.932	0.01	1.000	0.02	1.000
LEV	82	2	11	0.318	0.65	0.999	1.18	0.978
LEV-Collapsed	82	2	2	0.013	2.73	0.312	3.70	0.184
	DeJong and Ripoll (2006): panel with 60 countries, 5 periods							
DIF	20	8	64	0.257	0.81	1.000	73.4	< 0.001
DIF-Collapsed	200	8	28	0.397	0.63	0.995	0.87	1.000
LEV	260	8	28	0.830	0.39	1.000	0.69	1.000
LEV-Collapsed	260	8	8	0.107	0.22	0.999	0.32	1.000

Notes: We follow the original papers in utilizing heteroskedasticity-robust standard errors and associated degrees of freedom adjustments for the first-stage test statistics. All estimates are obtained using 2SLS. Details on the DIF- and LEV-(Collapsed) instrument matrices are provided in the text. In the LEV-Collapsed row for Voitchovsky (2005), we use the critical values for the bias test based on 3 instruments since 2 instrument critical values cannot be calculated for the case of two endogenous variables. See the notes to Table 3.1 for more details on the Kleibergen-Paap and Cragg-Donald tests, which apply in each regression to the full set of endogenous right-hand-side variables described in the text. †This study includes additional "external" instruments, log population and log area, which affect the diagnostic tests. Treating the "internal" GMM instruments as potentially weak and the external size instruments as likely strong (see Table 3.4), we apply the Hahn et al. (2011) test of the null hypothesis that the size instruments are valid and find p-values of 0.188 for the DIF equation, 0.228 for the DIF-Collapsed equation, < 0.001 for the LEV equation, and < 0.001 for the LEV-Collapsed equation. These results suggest that (i) log population provides a valid instrument in the difference equations but the difference in log population fails the exclusion restriction in the levels equations. Difference-in-Hansen tests, although not robust to weak instruments, yield similar insights.

3.8 Appendix

3.8.1 Background on the Rajan and Subramanian (2008) Instrument

Rajan and Subramanian (2008) construct an instrumental variable for the aid receipts in a "zero-stage" specification by regressing bilateral aid flows as a fraction of recipient GDP on recipient and donor characteristics. They use the resulting coefficients to calculate predicted bilateral aid flows. They sum these predicted bilateral flows across donors to arrive at predicted total aid receipts for each recipient country as a fraction of recipient GDP. This predicted total, a constructed instrument for true aid receipts, becomes the excluded instrument in a series of two-stage least squares regressions of economic growth on aid receipts and a set of control variables. The instrument is: $a_{dr} \equiv \frac{A_{dr}}{Y_r} = \sum_{j=0}^{7} \beta_j I_{i,dr} + \sum_{j=0}^{5} \beta_{i+8} (\ln P_d - \ln P_r) I_{j,dr} + \upsilon_{dr}$, where A_{dr} is dollars of aid given by donor d to recipient r, Y_r is the GDP of r, β_0 through β_{13} are regression coefficients, P_d is donor-country population, and P_r is recipient-country population. The I's are a set of time-invariant country dummy variables describing the country dyad: a current or past colonial relationship (I_1); a current or past colonial relationship with the United Kingdom (I_2), France (I_3), Spain (I_4), or Portugal (I_5); common language (I_6); and a current colonial relationship (I_7). Finally, $I_{0,dr} = 1 \; \forall \; d, r$ and υ_{dr} is an error term. The estimated coefficient vector $\widehat{\beta}$ is then used to generate predicted bilateral flows $\overline{a}_{dr}$, which are summed across donors to create the constructed instrument $\overline{a}_r = \sum_d \overline{a}_{dr}$, which then instruments for aid receipts $a_r \equiv A_r/Y_r$ in the cross-section growth regression $g_r = \gamma_1 a_r + \mathbf{X}_r' \mathbf{\Theta} + u_r$, where g_r is real GDP per capita growth, $\mathbf{X}_r$ is a vector of country characteristics, γ_1 is a regression coeffcient, Θ is a vector of regression coefficients, and u_r is an error term.

3.8.2 Testing for underidentification and weak instruments

We provide here additional details on the test statistics and inference procedures used in the paper to assess the strength of identification in regressions based on instrumental variables procedures. These weak instruments test statistics are often reported in empirical applications. However, the inferential implications, particularly for the weak instruments test statistics, are often left unstated.

The first diagnostic tool for assessing the strength of identification is based on a Langrange-Multiplier (LM) test for underidentification using the Kleibergen and Paap (2006) rk statistic. This test, readily implemented in Stata using the `ranktest` package, allows researchers to determine whether the minimal canonical correlation between the endogenous variables and the instruments is statistically different from zero. Another way of framing the test is by asking whether, after partialling out exogenous covariates and cross-correlations with the other endogenous variables and instruments, does the weakest correlation between an instrument and one of the endogenous variables suffice to contribute enough independent variation to add to the empirical rank of the instrument matrix? The p-values for this test are readily available after running the 2SLS estimation using the `ivreg2` package for Stata. The LM test for underidentification provides a lower hurdle than the tests for weak instruments.

The second set of diagnostics are based on the Stock and Yogo (2005) characterization of weak instruments using the first-stage F statistic and its multivariate analogue, the Cragg-Donald Wald statistic or its robust counterpart, the Kleibergen-Paap Wald statistic. The usual approach in the applied literature is to conclude that instruments are weak if these test statistics exceed the critical values tabulated by Stock and Yogo. Much less common is the full use of the testing procedures detailed in Section 4 of Stock and Yogo (2005), which provides richer probabilistic tools for characterizing weak instruments. Here, we provide a few practical details on how to construct p-values for the weak-instruments tests introduced in Section 3.3 and used throughout our paper. Adapting the empirical procedures in Gauss deployed in Yogo (2004), the formulation of p-values in Stata proceeds as follows:

1. Obtain the asymptotic threshold values for the concentration parameter Λ corresponding to the weak instruments test (relative OLS bias or t-test size). These values are contained in a number of Gauss matrices on Motohiro Yogo's

website.[40] We have converted these into Stata datasets (`lambfitBias.dta` and `lambfitSize.dta`) and provided them in supplementary material in the appendix.

2. Select the relevant value $\widehat{\Lambda}$ from the appropriate column and row of the lambfitBias or lambfitSize matrices based on the number of endogenous variables, the number of instruments K, and the level of bias or size distortion of interest. The relative OLS bias test is based on the finite sample distribution of the 2SLS estimator and hence critical values can only be calculated for cases where there are at least two more overidentifying restritions than the number of endogenous variables. In all specifications where this condition is not met, we report the weak instruments test based on the size distortion of the t-test. Critical values for this test, however, are not tabulated for cases with more than two endogenous variables. Thus, in cases with more endogenous variables and/or instruments than available in the Stock-Yogo tabulations, we take the penultimate available critical value in the given row and column of the table.

3. Obtain the Cragg-Donald ($\widehat{CD}$) and Kleibergen-Paap ($\widehat{KP}$) Wald test statistics after estimating the given 2SLS growth regression using `ivreg2` in Stata.

4. Calculate the p-value for the given null hypothesis using the formula: $p = 1 - \texttt{nchi2}(K, K \times \widehat{\Lambda}, K \times \widehat{CD})$, where `nchi2` is the noncentral χ^2 distribution with degrees of freedom K and noncentrality parameter $K \times \widehat{\Lambda}$. The p-value is valid for the Cragg-Donald statistic, which assumes homoskedastic error terms. While the $\widehat{KP}$ is robust to non-i.i.d. errors, its insertion in the p-value formula does not immediately follow since the Stock-Yogo diagnostics were not originally formulated for the non-i.i.d. case. Nevertheless, in characterizing weak instruments, we follow others in the literature and report $\widehat{CD}$ as well as $\widehat{KP}$ for each specification. Thus, while acknowledging that the p-values using the $\widehat{KP}$ statistic are not asymptotically correct, we report

[40] Available WWW: https://sites.google.com/site/motohiroyogo/home/publications/TestingWI_Programs.zip?attredirects=0.

it along with that for the $\widehat{CD}$ statistic for each of the given bias or size tests.

3.8.3 Weak-instrument robust inference

In Section 4.5, we employ the weak-instrument robust testing procedure of Kleibergen (2002) to examine 2SLS dynamic panel equations in levels and first differences. This procedure has been introduced as a higher power alternative to the Anderson-Rubin statistic. Here, we describe the steps for applying this method. Suppose that the dynamic panel growth equation is given by equation (3):

$$g_{i,t} = \beta \ln y_{i,t-1} + \mathbf{x}'_{i,t}\boldsymbol{\gamma}_1 + \widetilde{\mathbf{x}}'_{i,t}\boldsymbol{\gamma}_2 + \psi_i + \nu_{i,t}.$$

Suppose that $\boldsymbol{x}$ is a j-dimensional vector of endogenous growth determinants and $\widetilde{\boldsymbol{x}}$ is a k-dimensional vector of exogenous growth determinants including indicators for the period t. After constructing the appropriate instrument matrices for this equation in levels (LEV) and first differences (DIF), the method proceeds as follows:

1. Define the $j+1$ dimensional grid of possible values for the joint confidence region of β and $\boldsymbol{\gamma}_1$. In Figure 3, we restrict attention to a relatively narrow range of parameter values. Our principle was simply to start from the 2SLS point estimates and ensure that we chose a sufficiently wide range of values on both sides of that point estimate to encompass many values above and below zero. In the most general albeit infeasible case, one would want to examine the whole real line for each of the $j+1$ parameters. Lastly, one defines the increments over which to step along the range of values for a given parameter.

2. For the m-th $j+1$-tuple $(\beta^m, \boldsymbol{\gamma}_1^m)$, define $\widehat{g}_{i,t} = g_{i,t} - \beta^m \ln y_{i,t-1} - \mathbf{x}'_{i,t}\boldsymbol{\gamma}_1^m$ for the LEV equation and $\widehat{\Delta g}_{i,t} = \Delta g_{i,t} - \beta^m \Delta \ln y_{i,t-1} - \Delta \mathbf{x}'_{i,t}\boldsymbol{\gamma}_1^m$ for the DIF equation.

3. Regress $\beta^m \ln y_{i,t-1}$ $(\beta^m \Delta \ln y_{i,t-1})$ on all LEV (DIF) equation instruments and the exogenous covariates $(\Delta)\widetilde{\mathbf{x}}'_{i,t}$. Obtain the predicted values $\beta^m \widehat{\ln y_{i,t-1}}$

$(\beta^m \widehat{\Delta \ln y_{i,t-1}})$. Repeat the procedure for each of the j endogenous covariates in $\mathbf{x}'_{i,t}$.

4. Regress $\widehat{\Delta g_{i,t}}$ on $\beta^m \widehat{\Delta \ln y_{i,t-1}}$ and $\widehat{\Delta \mathbf{x}'_{i,t}\boldsymbol{\gamma}_1^m}$. Do the same for the LEV equation.

5. Test the joint significance of the right-hand side variables and store the associated p-values based on the large-sample $\chi^2(j+1)$ statistic for the given $j+1$-tuple $(\beta^m, \boldsymbol{\gamma}_1^m)$.

6. Using the resulting dataset comprised of p-values and $j+1$-tuples $(\beta^m, \boldsymbol{\gamma}_1^m)$ for the DIF and LEV equations, plot two-dimensional joint confidence ellipses (using the user-written `ellip` in Stata) for those values of $j+1$-tuples such that the p-value is greater than 0.05. More complex three-dimensional ellipsoids can be plotted in Matlab.

3.8.4 Weak identification of nonlinear effects in Rajan & Subramanian (2008)

If there are diminishing returns to capital in an economy, the effect of aid on growth can be nonlinear and concave. Assuming a linear relationship can easily cloud such a relationship: the best linear fit to a concave parabola has slope zero (presuming the full parabola is observed). Beyond this clear theoretical reason to test for nonlinear effects, several important aid-growth regressions published in the past decade have tested for and found a nonlinear relationship (e.g., Hansen and Tarp, 2001; Dalgaard et al., 2004). In a small part of one table, Rajan and Subramanian attempt to test for a nonlinear relationship between aid and growth, but their identification strategy does not allow this. The instrumentation in these regressions is extremely weak. They do not report this.

Columns 1, 4, and 7 of Table 3.9 show the underidentification and weak-instrument test statistics (p-values) for three regressions in Table 4 of Rajan and Subramanian (2008), where the aid effect is assumed linear. Instrumentation is strong. Columns 2, 5, and 8 show the same statistics for three regressions in their Table 7 (Panel A), which include a squared aid regressor, and use $\overline{a}_r$ and its square as the only excluded instruments. The inclusion of the squared term causes instrumentation strength to collapse in the periods 1980-2000 and 1990-2000, which is not reported in RS. Strength is retained in the 1970-2000 period, but solely due to the presence of Guinea-Bissau in the sample for that period (Guinea-Bissau is omitted from the sample in RS's other two periods). Without Guinea-Bissau, in columns 3, 6, and 9 , no useful degree of instrument strength is present regardless of periodization. In fact, we cannot reject that the structural equation is underidentified. All instrumentation in these nonlinear regressions, then, depends on a single country in a single period. The RS instrument does not allow a meaningful test of a nonlinear effect of aid on growth.[41]

There is no escape from this problem within the RS framework: The instruments independent of country size (I_1–I_7) do not explain aid variance, and the only strong instrument (population) is plausibly invalid, as we demonstrate in Section 3.3 of the paper. A more fruitful way forward is to find new instruments—better natural experiments to isolate the true effect of aid.

[41]One alternative procedure would be to carry out two separate zero-stage regressions, with regressands of linear aid and squared aid, to create two constructed instruments. This does not, however, improve instrument strength.

Table 3.9: Weak instruments in nonlinear specifications of Rajan and Subramanian (2008, Tables 4A and 7A)

Period	1970-2000			1980-2000			1990-2000		
Aid Specification	Linear	Quadratic	Quadratic sans GNB∓	Linear	Quadratic	Quadratic sans GNB∓	Linear	Quadratic	Quadratic sans GNB∓
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Kleibergen-Paap LM test p-value†	0.0004	<0.0001	0.397	0.0002	0.837	0.837	0.014	0.363	0.363
Cragg-Donald Waldstat‡	31.63	13.70	0.412	29.37	0.012	0.012	8.52	0.14	0.14
H_0: t-test size>10% (p-value)	0.058	0.085	0.999	0.085	1.000	1.000	0.871	1.000	1.000
H_0: t-test size>25% (p-value)	0.001	0.008	0.983	0.001	1.000	1.000	0.285	0.996	0.996
H_0: relative OLS bias>10% (p-value)	0.005	0.109	0.999	0.008	1.000	1.000	0.538	1.000	1.000
H_0: relative OLS bias>30% (p-value)	0.001	0.0210	0.993	0.001	1.000	1.000	0.275	0.998	0.998
Kleibergen-Paap Wald stat‡	36.12	13.10	0.279	31.26	0.017	0.017	6.952	0.314	0.314
H_0: t-test size>10% (p-value)	0.025	0.105	1.000	0.061	1.000	1.000	0.921	0.999	0.999
H_0: t-test size>25% (p-value)	0.0001	0.011	0.990	0.001	1.000	1.000	0.388	0.988	0.988
H_0: relative OLS bias>10% (p-value)	0.001	0.132	1.000	0.005	1.000	1.000	0.647	1.000	1.000
H_0: relative OLS bias>30% (p-value)	0.0001	0.028	0.996	0.001	1.000	1.000	0.376	0.995	0.995

Notes: The estimates in columns 1, 3 and 7 are exact replications of columns 2, 3, and 4 in Table 4A of Rajan and Subramanian (2008). The estimates in columns 2, 4 and 8 are exact replications of columns 2, 3, and 4 in Table 7A of Rajan and Subramanian (2008). ∓ Guinea-Bissau is only included in the 1970-2000 regressions in the original Table 7A. † The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (i.e., the rank condition fails). The test uses a rank test procedure from Kleibergen and Paap (2006). ‡ The Cragg-Donald and Kleibergen-Paap Wald statistics correspond respectively to non-robust and heteroskedasticity-robust multivariate analogues to the first-stage F statistics. Below each test statistic, we report the p-values from tests of whether (i) the actual size of the t-test(s) that $\beta_{aid} = 0$ (and $\beta_{aid^2} = 0$) at the 5% significance level is greater than 10 or 25%, and (ii) the bias of the IV estimates of β_{aid} (and β_{aid^2}) are greater than 10 or 30% of the OLS bias. In both cases, the critical values are obtained from Stock and Yogo (2005). Although critical values do not exist for the Kleibergen-Paap statistic, we follow the approach suggested in Baum et al. (2007) and apply the Stock and Yogo critical values initially tabulated for the Cragg-Donald statistic. The critical values for (ii) are (less conservatively) based on three instruments since one cannot calculate critical values in the (finite-sample)bias tests for the case of one endogenous variable and fewer than three instruments.

3.8.5 Further empirical and simulation results

Rajan & Subramanian Cross-Section Regressions (for 1990-2000)

Table 3.10 reproduces the Rajan and Subramanian (2008) results from Table 2 in the paper with an additional three columns covering their period 1990-2000. The results are similar to those for the longer periods (1970- and 1980-2000) as discussed in the paper.

Sources of identification in the Hausmann et al (2007) five-year panel

Using the Hausmann et al. (2007) panel data for the five-year periodization, Table 3.11 reports the same set of specification tests as in Table 4 based on their ten-year periodization. As noted in Section 3.3, the key result that export diversity (EXPY) increases growth does not hinge on the excludability of population size in the same restrictive manner that it did in the shorter panel with ten-year periodization. Although the result becomes null in column 6 when controlling for country size directly in the second stage, the effect of EXPY on growth is relatively robust to increasingly relaxing the excludability of the country size instruments in the levels and difference equation instrument matrices in columns 2-5. Nevertheless, there still remain concerns about the validity of the size instruments. While we cannot reject the validity of the size instruments on the basis of the difference-in-Hansen statistic in the specifications of columns 2-4, further unpacking of the levels and difference equation moment conditions in Section 4.4 revealed the validity of the size instruments could not be rejected for the levels equation according to the Hahn et al. (2011) test, the details of which are reported in the notes to Table 8.

Other measures of financial intermediation in Levine et al (2000)

Tables 3.12 and 3.13 estimate the same specifications as Table 6 using the two other measures of financial depth in Levine et al. (2000): private credit/GDP and the ratio of commercial to central bank credit, respectively. As noted in Section Section 4.2, the weak instruments problem of the system GMM estimator holds for these additional measures of financial depth. This can be seen most readily from the p-values for the weak-instruments tests reported in columns 6-9 of each table. The one slight difference with the liquid liabilities results in Table 6 is that we can reject the null of underidentification for the levels equation estimated using the collapsed instrument matrix in column 9 of Tables 3.12 and 3.13. Although these instruments pass the lower hurdle of underidentification, they remain weak.

Simulation results for a larger autoregressive parameter $\beta = 0.8$

Using the simulation procedure described in the paper, Figures 3.4 and 3.5 demonstrate the performance of the difference and system GMM estimators of γ (the coefficient on the endogenous growth determinant) when the persistence of the autoregressive parameter β increases from the baseline value of 0.2 to 0.8 (see equation (4)). The results are qualitatively unchanged from the baseline presented in Figures 1 and 2.

Table 3.10: Instrumentation strength in Rajan and Subramanian (2008) cross-section regressions

Period	1970–2000 ($N = 78$)			1980–2000 ($N = 75$)			1990–2000 ($N = 70$)		
"Zero-Stage" Specification	Replication	Colonial vars. only	Population vars. only	Replication	Colonial vars. only	Population vars. only	Replication	Colonial vars. only	Population vars. only
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Point estimate: Aid/GDP	0.096	–15.944	0.078	–0.004	–0.308	–0.028	–0.389	–0.035	–0.294
	(0.070)	(633.474)	(0.067)	(0.095)	(0.389)	(0.084)	(0.194)	(0.442)	(0.144)
CLR confidence set*: Aid/GDP	[–0.027,0.292]	$(-\infty, \infty)$	[–0.039,0.254]	[–0.186,0.232]	$(-\infty, \infty)$	[–0.194,0.170]	[–1.463,–0.063]	$(-\infty, 7.860] \cup [8.071, \infty)$	[–0.874,–0.021]
Kleibergen-Paap LM test (p-value)†	0.0004	0.978	0.0001	0.0002	0.282	0.0001	0.014	0.288	0.004
Cragg-Donald Wald stat‡	31.63	0.001	35.90	29.37	1.41	40.54	8.52	1.69	12.86
H_0: t-test size>10% (p-value)	< 0.001	0.999	< 0.001	0.001	0.888	< 0.001	0.303	0.865	0.118
H_0: t-test size>25% (p-value)	< 0.001	0.980	< 0.001	< 0.001	0.341	< 0.001	0.013	0.295	0.002
H_0: relative OLS bias>10% (p-value)	< 0.001	0.996	< 0.001	< 0.001	0.772	< 0.001	0.161	0.735	0.049
H_0: relative OLS bias>30% (p-value)	< 0.001	0.987	< 0.001	< 0.001	0.503	< 0.001	0.040	0.455	0.008
Kleibergen-Paap Wald stat‡	36.12	0.001	31.62	31.26	1.41	39.65	6.95	1.18	9.00
H_0: t-test size>10% (p-value)	< 0.001	0.999	< 0.001	0.001	0.888	< 0.001	0.407	0.906	0.275
H_0: t-test size>25% (p-value)	< 0.001	0.984	< 0.001	< 0.001	0.340	< 0.001	0.026	0.385	0.011
H_0: relative OLS bias>10% (p-value)	< 0.001	0.997	< 0.001	< 0.001	0.770	< 0.001	0.239	0.801	0.142
H_0: relative OLS bias>30% (p-value)	< 0.001	0.990	< 0.001	< 0.001	0.502	< 0.001	0.071	0.546	0.034

Notes: In all specifications, the instrumental variable is aid/GDP predicted from the zero-stage regression. The dependent variable in all specifications is average annual growth in GDP per capita over the period. Heteroskedasticity-robust standard errors in parentheses. Following the original paper, we retain the degrees-of-freedom adjustment to the Kleibergen-Paap F and LM statistics based on robust standard errors. For each of the three periods, the first column is based on exact replication of the baseline result in Rajan and Subramanian (2008, Table 4); the second column removes donor and recipient population terms from the zero-th stage specification used to estimate the predicted aid/GDP instrument $\overline{a}_r$, retaining only the colonial ties indicators; the third column retains only the population terms in the zero-th stage. All specifications include dummies for sub-Saharan Africa and East Asia. *The CLR confidence set corresponds to the weak-instrument robust confidence set obtained using the conditional likelihood ratio test in Moreira (2003). †The null hypothesis of the Kleibergen-Paap LM test is that the structural equation is underidentified (i.e., the rank condition fails). The test uses a rank test procedure from Kleibergen and Paap (2006). ‡In this special case of a single endogenous regressor, the Cragg-Donald and Kleibergen-Paap Wald statistics reduce respectively to the standard non-robust and heteroskedasticity-robust first-stage F statistics. Below each, we report the p-values from tests of whether (i) the actual size of the t-test that $\beta_{aid} = 0$ at the 5% significance level is greater than 10 or 25%, and (ii) the bias of the IV estimates of β_{aid} reported in the table are greater than 10 or 30% of the OLS bias. In both cases, the critical values are obtained from Stock and Yogo (2005). Although critical values do not exist for the Kleibergen-Paap statistic, we follow the approach suggested in Baum et al. (2007) and apply the Stock and Yogo critical values initially tabulated for the Cragg-Donald statistic. The critical values for (ii) are (less conservatively) based on three instruments since Stock and Yogo do not tabulate critical values in the bias tests for the case of one endogenous variable and fewer than three instruments.

Table 3.11: The (non-?)excludability of country size in 5-year panels of Hausmann et al. (2007)

Dependent variable	Growth	Growth	Growth	Growth	Growth	Growth
Estimator	IV$^{\mp}$	GMM-SYS$^{\mp}$	GMM-SYS	GMM-SYS	GMM-SYS	GMM-SYS
Size Instruments?	Yes	Yes	Yes, lev. Eq.	Yes, diff. eq.	No	Yes
Size Excluded?	Yes	Yes	Yes	Yes	Yes	No
	(1)	(2)	(3)	(4)	(5)	(6)
log initial GDP/capita	-0.030	-0.014	-0.015	-0.014	-0.008	-0.005
	(4.820)	(2.655)	(2.764)	(2.139)	(1.394)	(0.748)
log initial EXPY	**0.074**	**0.045**	**0.046**	**0.046**	**0.036**	**0.016**
	(5.105)	**(4.097)**	**(4.204)**	**(3.828)**	**(3.006)**	**(1.112)**
log human capital	0.004	0.004	0.003	-0.000	0.000	0.001
	(1.781)	(0.920)	(0.904)	(0.088)	(0.067)	(0.207)
log area						0.014
						(3.979)
log population						-0.009
						(3.233)
Observations	604	604	604	604	604	604
Number of Countries	79	79	79	79	79	79
Number of Periods	8	8	8	8	8	8
Number of Instruments	2	75	75	75	73	75
Hansen J test (p-value)	< 0.0001	0.507	0.502	0.467	0.623	0.267
Hansen J test excluding size instruments (p-value)		0.562	0.552	0.537	—	—
Difference-in-Hansen test or C statistic (p-value)$^{\pm}$		0.173	0.184	0.184	—	—
Kleibergen-Paap LM test (p-value)†	< 0.001	—	—	—	—	—
Cragg-Donald Waldstat‡	39.09	—	—	—	—	—
H_0: t-test size>25% (p-value)	< 0.001	—	—	—	—	—
H_0: relative OLS bias>30% (p-value)	< 0.001	—	—	—	—	—
Kleibergen-Paap Wald stat‡	34.25	—	—	—	—	—
H_0: t-test size>25% (p-value)	< 0.001	—	—	—	—	—
H_0: relative OLS bias>30% (p-value)	< 0.001	—	—	—	—	—

Notes: The dependent variable in all specifications is average annual growth over the period. The size instruments include log population and log area. The internal instruments refer to the lagged levels and lagged differences of endogenous right-hand side variables in the respective difference and levels equations of the dynamic panel GMM system of equations. $^{\mp}$ Columns 1 and 2 are based on Hausmann et al. (2007, Table 9, Columns 2 and 4). The null hypothesis of the difference-in-Hansen test (or C statistic, see Hayashi, 2000) is that the size instruments are valid. Following the original paper, we report heteroskedasticity-robust standard errors in parentheses and retain associated degrees of freedom adjustments for the first-stage test statistics. See the notes to Table 3.10 for more details on the Kleibergen-Paap and Cragg-Donald tests, which apply in column 1 to the endogenous EXPY variable.

Table 3.12: Weak instruments in panel regressions using private credit in Levine et al. (2000)

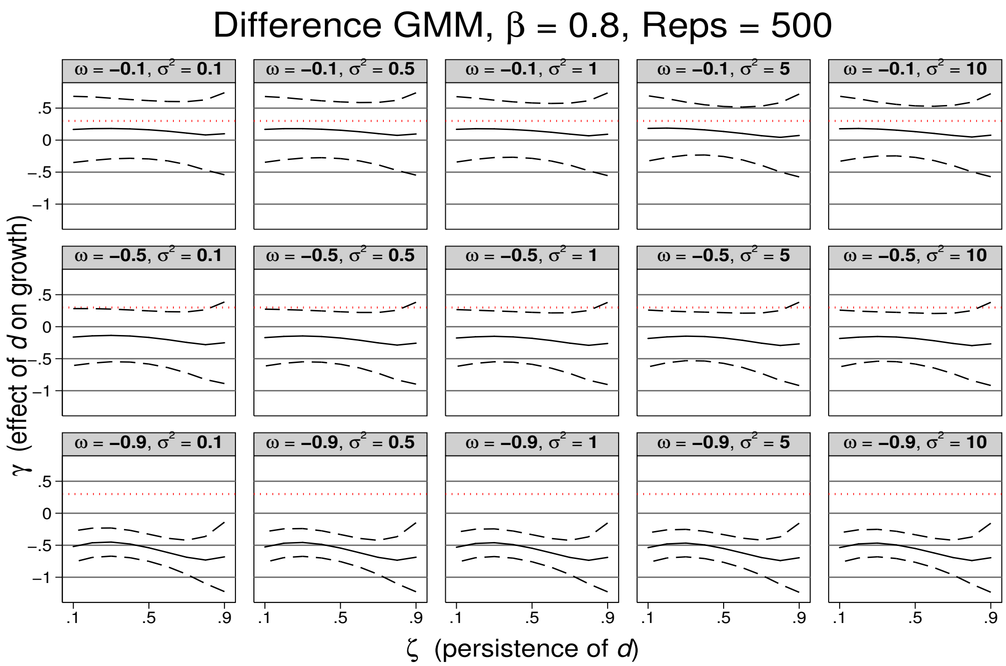
						Difference Equation		Levels Equation	
Estimator	GMM-SYS∓	GMM-SYS∓	OLS	OLS-FD	OLS-FE	2SLS	2SLS	2SLS	2SLS
Collapsed IV matrix	No	No	—	—	—	No	Yes	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Private credit	1.522	1.494	0.823	0.807	0.945	0.284	9.291	1.451	2.320
	(0.001)	(0.001)	(0.004)	(0.046)	(0.040)	(0.826)	(0.864)	(0.033)	(0.109)
Log initial GDP/capita	-0.364	-0.398	-0.315	-14.016	-7.832	-12.420	0.552	0.593	1.109
	(0.001)	(0.001)	(0.088)	(0.000)	(0.000)	(0.000)	(0.989)	(0.448)	(0.614)
					Other parameter estimates omitted				
N	359	359	345	323	345	323	323	345	345
Number of countries	74	74	74	74	74	74	74	74	74
Number of instruments	75	75	—	—	—	40	12	40	12
IV: Lagged levels	Yes	Yes	—	—	—	Yes	Yes	No	No
IV: Lagged differences	Yes	Yes	—	—	—	No	No	Yes	Yes
Kleibergen-Paap LM test (p-value)	—	—	—	—	—	0.249	0.879	0.635	0.069
Cragg-Donald Waldstat	—	—	—	—	—	0.73	0.004	0.67	0.51
H_0: relative OLS bias>10% (p-value)	—	—	—	—	—	1.000	1.000	1.000	1.000
H_0: relative OLS bias>30% (p-value)	—	—	—	—	—	1.000	1.000	1.000	0.987
Kleibergen-Paap Wald stat	—	—	—	—	—	1.08	0.001	1.16	0.47
H_0: relative OLS bias>10% (p-value)	—	—	—	—	—	1.000	1.000	1.000	1.000
H_0: relative OLS bias>30% (p-value)	—	—	—	—	—	1.000	1.000	1.000	0.983

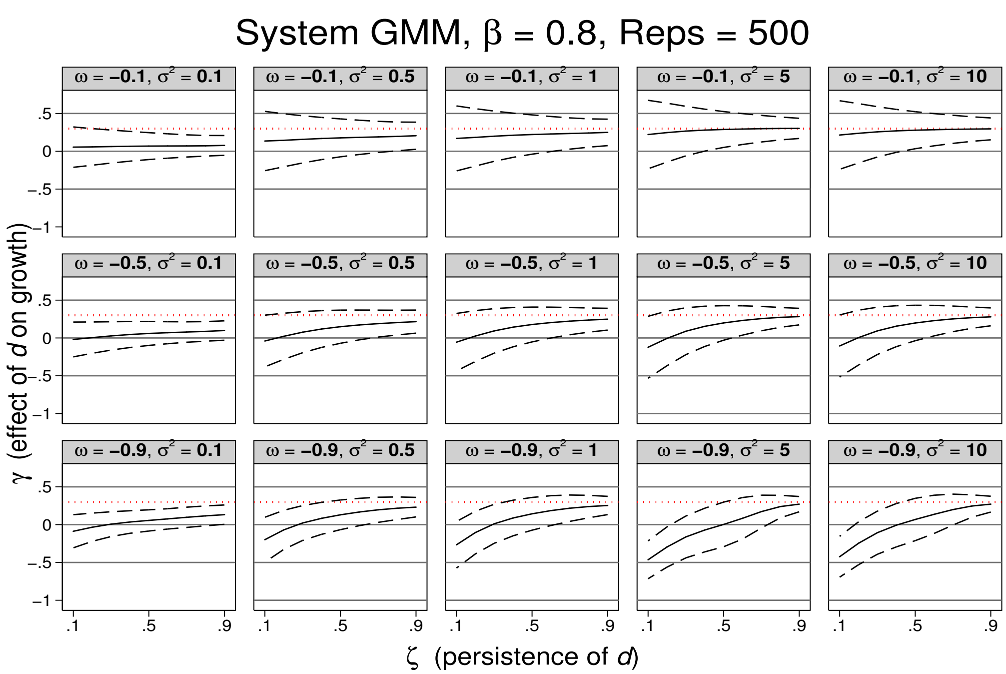
Notes: The dependent variable in all specifications is average annual growth in GDP per capita each period. ∓ Column 1 reproduces the published version of Levine et al. (2000, Table 5, Column 2), and column 2 reports our best attempted replication using the `DPD96` program for Gauss, the publicly available dataset, and a Gauss program used to generate their results provided by Thorsten Beck. Further details on the difference in sample sizes across columns, our replication efforts, and the associated differences in the Gauss and Stata programs for dynamic panel GMM regressions can be found in Appendix 3.8.6. The following variables are included in the regressions but suppressed in the table here for presentational purposes: government size, openness to trade, inflation, average years of secondary schooling, black market premium, time period dummies and a constant. The first five of these variables are treated as endogenous. Following the original paper, we report p-values in parentheses. See the notes to Table 3.10 for more details on the Kleibergen-Paap and Cragg-Donald tests, which apply in columns 6-9 to the full set of endogenous right-hand-side variables.

Table 3.13: Weak instruments in panel regressions using commercial vs. central bank credit in Levine et al. (2000)

						Difference Equation		Levels Equation	
Estimator	GMM-SYS∓	GMM-SYS∓	OLS	OLS-FD	OLS-FE	2SLS	2SLS	2SLS	2SLS
Collapsed IV matrix	No	No	—	—	—	No	Yes	No	Yes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Commercial vs. Central Bank Credit	2.437	2.293	1.243	2.526	3.397	3.649	-7.670	0.864	4.708
	(0.001)	(0.001)	(0.033)	(0.017)	(0.010)	(0.158)	(0.979)	(0.536)	(0.198)
Log initial GDP/capita	-0.117	-0.348	-0.138	-13.896	-7.912	-12.817	-8.985	0.568	1.470
	(0.223)	(0.015)	(0.419)	(0.000)	(0.000)	(0.000)	(0.963)	(0.347)	(0.483)
	Other parameter estimates omitted								
N	359	359	345	324	345	324	324	345	345
Number of countries	74	74	74	74	74	74	74	74	74
Number of instruments	75	75	—	—	—	40	12	40	12
IV: Lagged levels	Yes	Yes	—	—	—	Yes	Yes	No	No
IV: Lagged differences	Yes	Yes	—	—	—	No	No	Yes	Yes
Kleibergen-Paap LM test (p-value)	—	—	—	—	—	0.489	0.963	0.500	0.022
Cragg-Donald Waldstat	—	—	—	—	—	0.71	< 0.001	0.88	0.76
H_0: relative OLS bias>10% (p-value)	—	—	—	—	—	1.000	1.000	1.000	1.000
H_0: relative OLS bias>30% (p-value)	—	—	—	—	—	1.000	1.000	1.000	0.954
Kleibergen-Paap Wald stat	—	—	—	—	—	0.84	< 0.001	1.25	0.75
H_0: relative OLS bias>10% (p-value)	—	—	—	—	—	1.000	1.000	1.000	1.000
H_0: relative OLS bias>30% (p-value)	—	—	—	—	—	1.000	1.000	1.000	0.954

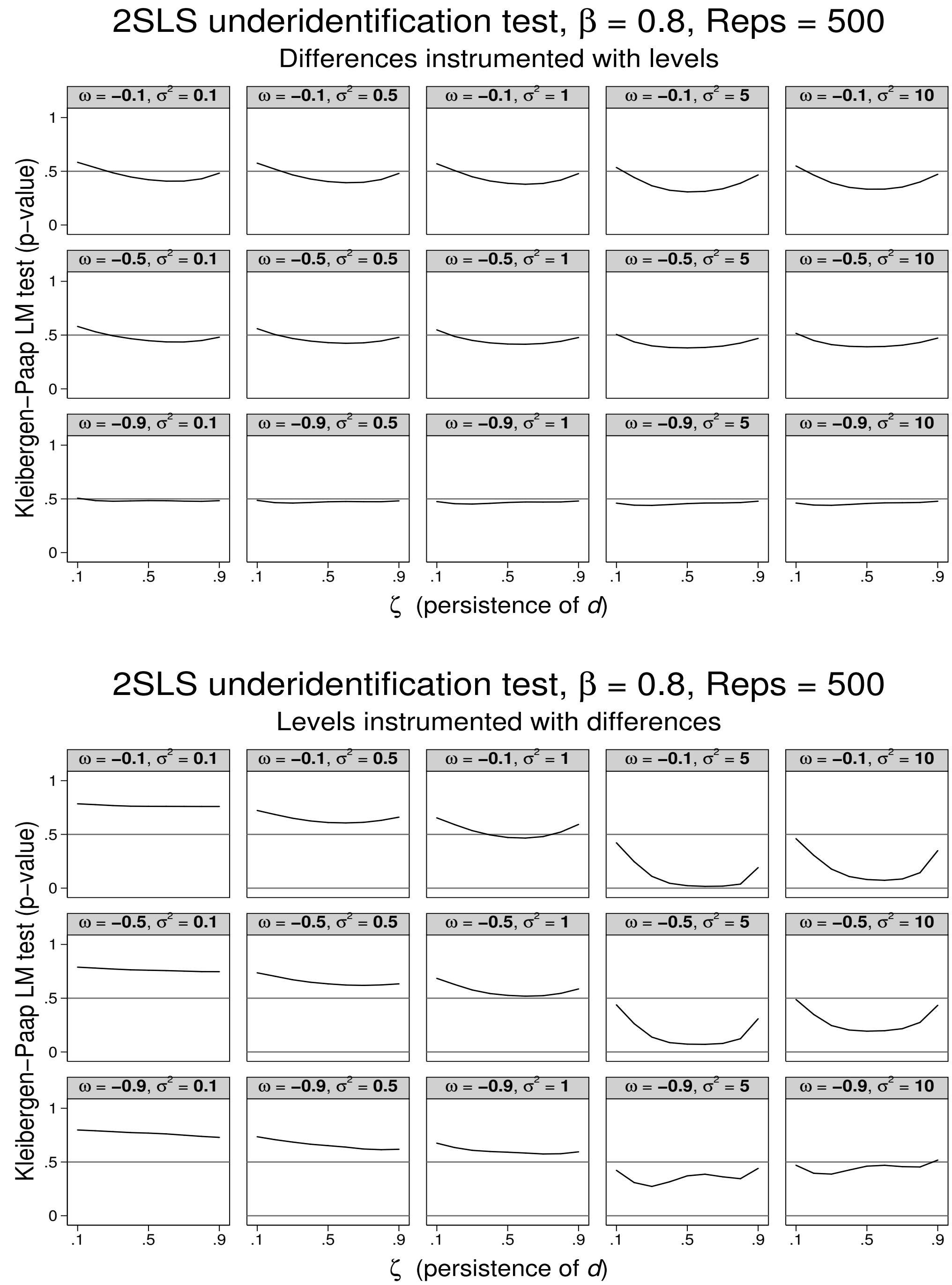
Notes: The dependent variable in all specifications is average annual growth in GDP per capita each period. ∓ Column 1 reproduces the published version of Levine et al. (2000, Table 5, Column 3), and column 2 reports our best attempted replication using the `DPD96` program for Gauss, the publicly available dataset, and a Gauss program used to generate their results provided by Thorsten Beck. Further details on the difference in sample sizes across columns, our replication efforts, and the associated differences in the Gauss and Stata programs for dynamic panel GMM regressions can be found in Appendix 3.8.6. The following variables are included in the regressions but suppressed in the table here for presentational purposes: government size, openness to trade, inflation, average years of secondary schooling, black market premium, time period dummies and a constant. The first five of these variables are treated as endogenous. Following the original paper, we report p-values in parentheses. See the notes to Table 3.10 for more details on the Kleibergen-Paap and Cragg-Donald tests, which apply in columns 6-9 to the full set of endogenous right-hand-side variables.





Notes: The graphs show parameter estimates and 95% confidence intervals from simulations of the model in equation (4) of the paper based on 500 draws of a sample size of 600 with 100 cross-sectional units and 6 time periods, fixed $\beta = 0.8$, varying $\zeta \in \{0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9\}$, varying degrees of endogeneity $\omega \in \{-0.1, -0.5, -0.9\}$, and alternative variances of the idiosyncratic shock, $\sigma^2 \in \{0.1, 0.5, 1, 5, 10\}$, where the variance of cross-sectional heterogeneity is fixed at 1. The dashed red line shows the true value of $\gamma = 0.3$ in the simulations.

Figure 3.4: Power and size properties of GMM estimators in simulation results



Notes: The graphs show p-values from a Kleibergen-Paap LM test for (the null of) underidentification in the levels and differences equations from simulations of the model in equation (4) in the paper as detailed in the notes to Figure 3.4. See the notes to Table 3.10 for details on the Kleibergen-Paap test.

Figure 3.5: Weak identification in simulation results, $\beta = 0.8$

3.8.6 Replicating growth studies

We describe here our replications of the empirical growth studies assessed in Sections 3 and 4.

Levine et al (2000)

Despite the provision by Levine et al. (hereafter, LLB) of a publicly available dataset (Financial_Intermediation_and_Growth_dataset.xls) on a World Bank website (http://go.worldbank.org/40TPPEYOC0), we faced a few difficulties in obtaining an exact replication of their dynamic panel GMM results. Nevertheless, on the basis of our replications efforts described here, we are highly confident that the subsequent OLS and 2SLS results that we report in Tables 6, 3.12, and 3.13 are those that LLB would have gotten at the time they wrote, with precisely the same data.

In the process of attempting to replicate the original LLB results using exactly the same version of their estimator in Gauss (`DPD96`), the same data, and the same program file provided by one of the LLB authors (Thorsten Beck), we discovered a bug in `DPD96`.[42] The bug produced different two-step GMM estimates across consecutive runs of the same program over the same data, even after reloading the data anew at each run. The result holds for the other two measures as well. While the estimates do not vary wildly, we believe that this sort of non-deterministic potential within this program for the deterministic dynamic panel GMM estimator could explain why the LLB result cannot be reproduced exactly within Gauss (or Stata).

Table 3.14 below compares the published parameter estimates in Table 5 of LLB to replications using the original data and the `DPD96` program in Gauss and the `xtabond2` program in Stata. Columns 2, 5, and 8 correspond to the estimates in column 2 of Tables 6, 3.12, and 3.13, respectively. The replication based on `DPD96` is quite close to the original published estimates. In only one instance does the sign of the parameter estimate differ (inflation for the private credit outcome).

[42] Before proceeding to the replication, we removed three countries from the excel dataset, which were not listed as part of the 74 country panel in Table 9 of their published paper.

Turning to the Stata replications in columns 3, 6, and 9, we find larger differences with the estimates obtained using Gauss despite setting all options in `xtabond2` to mimic the `DPD96` formulation (see Roodman, 2009a). Roodman (2009b) reports similar difficulties replicating their results.[43] The other point to notice is that the sample size apparently differs in the Gauss and Stata replications. This is actually not accurate, though. After inspection of the sample countries and years used in each, we find that the samples are identical and that `DPD96` output does not seem to be reporting the actual sample size.

Rajan & Subramanian (2008)

The original Rajan and Subramanian dataset and code were kindly provided by the authors. As noted in the paper, we exactly replicate their cross-section and dynamic panel results relevant to our discussion. The analysis in Tables 1, 2, and 5 meanwhile required us to supplement their original dataset with population data. The original dataset contained population ratios from zero-stage regressions but not separate figures for period-initial receiving country population. For the zero-stage regressions, the only database with sufficiently complete country coverage was the International Monetary Fund's online *International Financial Statistics* (accessed Sept. 9, 2007), which had populations of all aid recipient countries in the Rajan and Subramanian dataset, except for Bermuda, Kiribati, Turkmenistan, and Uzbekistan, which come from the World Bank's *World Development Indicators 2007*. In the main regressions, the extreme breadth of country coverage is not needed and we took population from the *Penn World Table 6.1*, since real GDP/capita came from that source. The correlation between the two sources' population estimates is near unity.

In their dynamic panel GMM results, Rajan and Subramanian include the

[43] Our initial efforts at replication were done in consultation with Roodman. Subsequently, after correspondence with Thorsten Beck, we obtained additional input into the Gauss replication. Our Stata replication for private credit slightly differs from that in Roodman (2009b) for two reasons. First, we do not use the Windmeijer (2005) two-step variance correction since this procedure was not available to LLB at the time of their study in the late 1990s. Second, we drop three countries from the publicly available excel dataset, which were not listed among the 74 countries in Table 9.

second through seventh lags as instruments for the difference equation in both specifications. They note that they are employing up to eight lags, but given that their panel consists of eight periods and only four of the five year periods since 1985 are actually used due to missing data on their institutional quality measure, their specifications naturally do not include eighth lagged levels as instruments for any of the endogenous regressors. Also, although they claim to include an additional set of time-invariant, excluded instruments in their main difference-equation specifications (geography, ethnic fractionalization, Sub-Saharan Africa and East Africa), a Stata coding error results in their being dropped from the equations regressing differenced endogenous variables on lagged levels. In Table 7 of the paper, to be consistent with their published results, we exclude these four time-invariant dummies from the Arellano-Bond regression in column 1 and the difference equation in the Blundell-Bond regression in column 2, as well as the corresponding 2SLS regressions in subsequent columns.

Hausmann et al (2007)

The original Hausmann et al. dataset and code were kindly provided by the authors. In Table 3.11, we exactly replicate their original pooled 2SLS and system GMM estimates for their panel based on a five-year periodization. In Table 4, despite applying their original code to the original data, we obtain slightly different estimates from those reported in their published paper for the system GMM specification on the panel with ten-year periodization. The pooled 2SLS estimates are identical. Nevertheless, the differences are trivial and in no way affect our main message in Table 4 (or the key findings in Hausmann et al.'s original paper for that matter).

DeJong & Ripoll (2006)

The original DeJong and Ripoll dataset and code were kindly provided by the authors. We are able to obtain exact replications of their dynamic panel GMM estimates in Table 2.

Hauk & Wacziarg (2009)

The original Hauk and Wacziarg dataset and code were kindly provided by the authors. We are able to obtain exact replications of their dynamic panel GMM estimates in Table 13.

Voitchovsky (2005)

The original Voitchovsky dataset was kindly provided by the author. Using the `DPD98` package (the successor to `DPD96`) for Gauss as originally deployed by the author, we are able to obtain a close replication of the system GMM estimates reported in Table 2 of the published paper. We could not obtain an exact replication of the published results likely due to the bug in the `DPD96` program noted above and inherited by the `DPD98`.

Voitchovsky (2005) constructs a non-standard set of instruments, motivated by arguments against using all the conventional Blundell-Bond moment conditions. For the DIF equation, the instruments include twice and thrice lagged income per capita, lagged investment, the twice lagged and difference in schooling rates, and the twice lagged difference in inequality measures. Anderson and Hsiao (1982) were the first to suggest using twice lagged differences as instruments for the lagged, differenced dependent variable in a dynamic panel setting (see also Arellano, 1989), though the typical Arellano and Bond (1991) or Blundell and Bond (1998) applications instrument contemporaneous differences with lagged levels, retaining the first lagged difference as an instrument for contemporaneous levels. For the LEV equation, the instruments include once lagged and differenced investment and schooling rates; the inequality measures in levels and lagged income per capita are treated (rather unconventionally) as exogenous in the levels equation.

Table 3.14: Replicating Levine et al. (2000)

	Original	Replication	Replication	Original	Replication	Replication	Original	Replication	Replication
SYS-GMM Estimator	`DPD96`	`DPD96`	`xtabond2`	`DPD96`	`DPD96`	`xtabond2`	`DPD96`	`DPD96`	`xtabond2`
Collapsed IV matrix	No	No	No	No	No	No	No	No	No
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Liquid liabilities	2.952	2.834	3.176						
	(0.001)	(0.001)	(0.000)						
Private credit				1.522	1.494	1.451			
				(0.001)	(0.001)	(0.000)			
Commercial vs. Central Bank Credit							2.437	2.293	1.383
							(0.001)	(0.001)	(0.000)
Log initial GDP/capita	-0.742	-0.792	-0.525	-0.364	-0.398	-0.268	-0.117	-0.348	-0.225
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.020)	(0.223)	(0.015)	(0.008)
Government size	-1.341	-1.419	0.249	-1.987	-1.841	-0.195	-1.13	-1.088	0.555
	(0.001)	(0.001)	(0.481)	(0.001)	(0.001)	(0.576)	(0.001)	(0.001)	(0.000)
Openness to trade	0.325	0.372	-0.047	0.442	0.499	-0.016	0.497	0.620	0.646
	(0.169)	(0.124)	(0.847)	(0.010)	(0.021)	(0.929)	(0.002)	(0.001)	(0.000)
Inflation	1.748	1.675	1.074	-0.178	0.055	-0.598	-1.772	-2.413	-1.802
	(0.001)	(0.001)	(0.000)	(0.543)	(0.810)	(0.007)	(0.001)	(0.001)	(0.000)
Avg. yrs. secondary school	0.780	0.732	0.041	0.639	0.472	0.195	0.638	0.775	0.935
	(0.001)	(0.001)	(0.743)	(0.001)	(0.001)	(0.128)	(0.001)	(0.001)	(0.000)
Black market premium	-2.076	-2.014	-2.102	-1.027	-1.109	-1.062	-1.044	-1.395	-1.018
	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)	(0.001)	(0.001)	(0.000)
Constant	0.06	1.061	-4.301	4.239	4.042	1.300	-5.677	-4.001	-5.713
	(0.954)	(0.195)	(0.002)	(0.001)	(0.001)	(0.119)	(0.001)	(0.001)	(0.277)
Observations	359	359	345	359	359	345	359	359	345
Number of countries	74	74	74	74	74	74	74	74	74
Number of instruments	75	75	75	75	75	75	75	75	75

Notes: The dependent variable in all specifications is average annual growth in GDP per capita each period. The following variables are included in the regressions but suppressed in the table here for presentational purposes: government size, openness to trade, inflation, average years of secondary schooling, black market premium, time period dummies and a constant. The first five of these variables are treated as endogenous. Following the original paper, we report p-values in parentheses based on two-step estimates without the Windmeijer (2005) correction, which became available after the LLB study.

Chapter 3, in part, has been published as it appears in American Economic Journal: Macroeconomics, 2012. Bazzi, Samuel. The dissertation author was the primary investigator and author of this paper.

Bibliography

Abadie, A., "Semiparametric Difference-in-Differences Estimators," *Review of Economic Studies*, 2005, *72* (1), 1–19.

_ **and G.W. Imbens**, "Large sample properties of matching estimators for average treatment effects," *Econometrica*, 2005, *74* (1), 235–267.

Abramitzky, R., L.P. Boustan, and K. Eriksson, "Have the poor always been less likely to migrate? Evidence from inheritance practices during the Age of Mass Migration," *NBER Working Paper*, 2012, *No. 18298.*

Acemoğlu, Daron, "Theory, General Equilibrium, Political Economy and Empirics in Development Economics," *Journal of Economic Perspectives*, 2010, *24* (2), 17–32.

_ **, Simon Johnson, and James A. Robinson**, "The Colonial Origins of Comparative: An Empirical Investigation," *American Economic Review*, 2001, *91* (5), 1370–1400.

_ **,** _ **,** _ **, and Pierre Yared**, "From Education to Democracy?," *American Economic Review Papers and Proceedings*, 2005, *95* (2), 44–49.

Albouy, David Y., "The Colonial Origins of Comparative Development: An Empirical Investigation: Comment," *American Economic Review*, forthcoming.

Alesina, Alberto and Romain Wacziarg, "Openness, Country Size, and Government," *Journal of Public Economics*, 1998, *69* (1), 305–321.

Alfaro, Laura, Areendam Chanda, Şebnem Kalemli-Özcan, and Selin Sayek, "FDI and Economic Growth: The Role of Local Financial Markets," *Journal of International Economics*, 2004, *64* (1), 89–112.

Anderson, J.E. and E. van Wincoop, "Trade Costs," *Journal of Economic Literature*, 2004, *42*, 691–751.

Anderson, T. W. and Cheng Hsiao, "Formulation and Estimation of Dynamic Models Using Panel Data," *Journal of Econometrics*, 1982, *18* (1), 47–82.

Anderson, T.W. and H. Rubin, "Estimation of the parameters of a single equation in a complete system of stochastic equations," *The Annals of Mathematical Statistics*, 1949, *20*, 46–63.

Angeles, Luis and Kyriakos C. Neanidis, "Aid Effectiveness: The Role of the Local Elite," *Journal of Development Economics*, 2009, *90* (1), 120–134.

Angelucci, M., "Migration and Financial Constraints: Evidence from Mexico," *Unpublished manuscript*, 2012.

_ **and O. Attanasio**, "The Demand for Food of Poor Urban Mexican Households: Understanding Policy Impacts Using Structural Models," *American Economic Journal: Economic Policy*, forthcoming.

Ardington, C., A. Case, and V. Hosegood, "Labor Supply Responses to Large Social Transfers: Longitudinal Evidence from South Africa," *American Economic Journal: Applied Economics*, 2009, *1*, 22–48.

Arellano, M., "A note on the Anderson-Hsiao estimator for panel data," *Economics Letters*, 1989, *31* (4), 337–341.

Arellano, Manuel and Olympia Bover, "Another Look At the Instrumental Variable Estimation of Error-Components Models," *Journal of Econometrics*, 1995, *68* (1), 29–52.

_ **and Stephen Bond**, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," *Review of Economic Studies*, 1991, *58* (2), 277–297.

Armenter, R. and M. Koren, "A Balls-and-Bins Model of Trade," *Unpublished manuscript*, 2010.

Ashley, Richard, "Assessing the Credibility of Instrumental Variables Inference with Imperfect Instruments Via Sensitivity Analysis," *Journal of Applied Econometrics*, 2009, *24* (2), 325–337.

Attanasio, O.P., C. Meghir, and A. Santiago, "Education choices in Mexico: Using a structural model and a randomized experiment to evaluate Progresa," *The Review of Economic Studies*, 2012, *79* (1), 37–66.

Bachtiar, P. B., "The Governance of Indonesian Overseas Employment in the Context of Decentralization," *SMERU Research Institute Working Paper*, 2011.

Baird, S., C. McIntosh, and B. Özler, "Cash or Condition? Evidence from a Cash Transfer Experiment," *The Quarterly Journal of Economics*, 2011, *126* (4), 1709–1753.

Baltagi, Badi H. and Young-Jae Chang, "Incomplete panels: A comparative study of alternative estimators for the unbalanced one-way error component regression model," *Journal of Econometrics*, 1994, *62* (1), 67–89.

Bank Indonesia, "Survei Nasional Pola Remitansi TKI," *Special Report*, 2009.

Barrera-Osorio, F., M. Bertrand, L.L. Linden, and F. Perez-Calle, "Improving the Design of Conditional Transfer Programs: Evidence from a Randomized Education Experiment in Colombia," *American Economic Journal: Applied Economics*, 2011, *3* (2), 167–195.

Barro, Robert J., "Economic Growth in a Cross Section of Countries," *Quarterly Journal of Economics*, 1991, *106* (2), 407–443.

_ , "Democracy and Growth," *Journal of Economic Growth*, 1996, *1* (1), 1–27.

Baum, Christopher F., Mark E. Schaffer, and Stephen Stillman, "Enhanced Routines for Instrumental Variables/GMM Estimation and Testing," *Stata Journal*, 2007, *7* (4), 465–506.

Baumol, William J., "Productivity Growth, Convergence, and Welfare: What the Long-run Data Show," *American Economic Review*, 1986, *76* (5), 1072–1085.

Bazzi, S., "International Migration from Indonesia: New Stylized Facts," *Unpublished manuscript*, 2012a.

_ , "A Note on Heterogeneity and Aggregation Using Data on Agricultural Productivity," *Unpublished manuscript*, 2012b.

_ **and C. Blattman**, "Economic Shocks and Conflict: The (Absence of) Evidence from Commodity Prices," *Unpublished Manuscript*, 2012.

_ **, S. Sumarto, and A. Suryahadi**, "It's All in the Timing: Evaluating the Expenditure and Labor Supply Response to Unconditional Cash Transfers," *Unpublished manuscript*, 2012.

Beck, Thorsten and Ross Levine, "Stock markets, banks, and growth: Panel evidence," *Journal of Banking & Finance*, 2004, *28* (3), 423–442.

_ **, Aslı Demirgüç-Kunt, and Ross Levine**, "SMEs, Growth, and Poverty: Cross-country Evidence," *Journal of Economic Growth*, 2005, *10* (3), 199–229.

Beine, M., F. Docquier, and C. Özden, "Diasporas," *Journal of Development Economics*, 2011.

Benhabib, J. and B. Jovanovic, "Optimal Migration: A World Perspective," *International Economic Review*, 2012, *53*, 321–348.

Benjamin, D., "Can unobserved land quality explain the inverse productivity relationship?," *Journal of Development Economics*, 1995, *46*, 51–84.

Bertoli, S., J. Fernández-Huertas Moraga, and F. Ortega, "Crossing the Border: Self-selection, Earnings and Individual Migration Decisions," *CREAM Discussion Paper*, 2010, *11/10.*

Bianchi, M. and M. Bobba, "Liquidity, Risk, and Occupational Choices," *Review of Economic Studies*, forthcoming.

Blundell, R., A. Bozio, and G. Laroque, "Labor Supply and the Extensive Margin," *American Economic Review*, 2011, *101*, 482–486.

_ and T.M. Stoker, "Heterogeneity and aggregation," *Journal of Economic Literature*, 2005, *43* (2), 347–391.

_ and _ , "Models of aggregate economic relationships that account for heterogeneity," *Handbook of Econometrics*, 2007, *6*, 4609–4666.

_ , L. Pistaferri, and I. Preston, "Consumption inequality and partial insurance," *American Economic Review*, 2008, *98* (5), 1887–1921.

Blundell, Richard and Stephen Bond, "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models," *Journal of Econometrics*, 1998, *87* (1), 115–143.

_ and _ , "GMM Estimation with Persistent Panel Data: An Application to Production Functions," *Econometric Reviews*, 2000, *19* (3), 321–340.

Blundell, Richard W., Stephen R. Bond, and Frank Windmeijer, "Estimation in dynamic panel data models: improving on the performance of the standard GMM estimator," *Advances in Econometrics*, 2000, *15*, 53–92.

Bobba, Matteo and Decio Coviello, "Weak Instruments and Weak Identification in Estimating the Effects of Education on Democracy," *Economics Letters*, 2007, *96* (2), 301–306.

Bond, Stephen R., Anke Hoeffler, and Jonathan Temple, "GMM Estimation of Empirical Growth Models," *CEPR Discussion Paper No. 3048.*, 2001.

Boone, Peter, "Politics and the Effectiveness of Foreign Aid," *European Economic Review*, 1996, *40* (2), 289–329.

Borensztein, Eduardo, Jose De Gregorio, and Jong-Wha Lee, "How Does Foreign Direct Investment Affect Economic Growth?," *Journal of International Economics*, 1998, *45* (1), 115–135.

Borger, S., "Self-Selection and Liquidity Constraints in Different Migration Cost Regimes," *Unpublished manuscript*, 2010.

Borjas, G., "Self-selection and the Earnings of Immigrants," *American Economic Review*, 1987, *77*, 531–553.

Bosworth, Barry P. and Susan M. Collins, "The Empirics of Growth: An Update," *Brookings Papers on Economic Activity*, 2003, *2*, 113–206.

Bowsher, Clive G., "On testing overidentifying restrictions in dynamic panel data models," *Economics Letters*, 2002, *77* (2), 211–220.

Brock, William A. and Steven N. Durlauf, "Growth Empirics and Reality," *World Bank Economic Review*, 2001, *15* (2), 229–272.

Broda, C. and J. A. Parker, "The Economic Stimulus Payments of 2008 and the Aggregate Demand for Consumption," *Unpublished Mansucript*, 2012.

Browning, M. and M. D. Collado, "The response of expenditures to anticipated income changes: panel data estimates," *American Economic Review*, 2001, *91* (3), 681–692.

Bryan, G., S. Chowdhury, and A. M. Mobarak, "Seasonal Migration and Risk Aversion," *Unpublished Manuscript*, 2012.

Bun, Maurice and Monique de Haan, "Weak instruments and the first stage F statistic in IV models with a nonscalar error covariance structure," *University of Amsterdam, Working Paper 2*, 2010.

Bun, Maurice J. G. and Frank Windmeijer, "The Weak Instrument Problem of the System GMM Estimator in Dynamic Panel Data Models," *Econometrics Journal*, 2010, *13* (1), 95–126.

Burnside, Craig and David Dollar, "Aid, Policies, and Growth," *American Economic Review*, 2000, *90* (4), 847–868.

Busso, M., J. DiNardo, and J. McCrary, "New Evidence on the Finite Sample Properties of Propensity Score Matching and Reweighting Estimators," *IZA Discussion Papers*, 2009.

Calderón, César, Alberto Chong, and Norman Loayza, "Determinants of current account deficits in developing countries," *Contributions to Macroeconomics*, 2002, *2* (1), Article 2.

Cameron, A. C., J. B. Gelbach, and D. L. Miller, "Bootstrap-based Improvements for Inference with Clustered Errors," *Review of Economics and Statistics*, 2008, *90*, 414–427.

_ , _ , and _ , "Bootstrap-based improvements for inference with clustered errors," *The Review of Economics and Statistics*, 2008, *90* (3), 414–427.

_ , _ , and _ , "Robust inference with multi-way clustering," *Journal of Business and Economic Statistics*, 2011, *28*, 238–249.

Cameron, L. and M. Shah, "Can Mistargeting Destroy Social Capital and Stimulate Crime? Evidence from a Cash Transfer Program in Indonesia," *Unpublished Manuscript*, 2012.

Carrillo, P.E. and J. Ponce Jarrín, "Efficient delivery of subsidies to the poor: Improving the design of a cash transfer program in Ecuador," *Journal of Development Economics*, 2009, *90* (2), 276–284.

Carrington, W. J., E. Detragiache, and T. Vishwanath, "Migration with Endogenous Moving Costs," *American Economic Review*, 1996, *86*, 909–930.

Caselli, Francesco, Gerardo Esquivel, and Fernando Lefort, "Reopening the Convergence Debate: A New Look At Cross Country Growth Empirics," *Journal of Economic Growth*, 1996, *1* (3), 363–389.

Cattaneo, M.D., "Efficient semiparametric estimation of multi-valued treatment effects under ignorability," *Journal of Econometrics*, 2010, *155* (2), 138–154.

Chambers, M.J. and R.E. Bailey, "A theory of commodity price fluctuations," *Journal of Political Economy*, 1996, pp. 924–957.

Clauset, A., C. R. Shalizi, and M. E. J. Newman, "Power-Law Distributions in Empirical Data," *Society for Industrial and Applied Mathematics*, 2009, *51*, 661–703.

Clemens, M. A., "Economics and Emigration: Trillion-Dollar Bills on the Sidewalk," *Journal of Economic Perspectives*, 2011, *25*, 83–106.

Clemens, Michael A., Steven Radelet, Rikhil R. Bhavnani, and Samuel Bazzi, "Counting Chickens When They Hatch: Timing And The Effects Of Aid On Growth," *The Economic Journal*, 2012, *122*, 590–617.

Coady, D., J. Tyson, J.M. Piotrowski, R. Gillingham, R. Ossowski, and S. Tareq, *Petroleum Product Subsidies: Costly, Inequitable, and On the Rise*, International Monetary Fund, 2010.

Conley, T.G., C.B. Hansen, and P.E. Rossi, "Plausibly exogenous," *The Review of Economics and Statistics*, forthcoming.

Conley, Timothy G., Christian B. Hansen, and Peter E. Rossi, "Plausibly Exogenous," *Review of Economics and Statistics*, 2012, *94* (1), 260–272.

Cragg, John G. and Stephen G. Donald, "Testing Identifiability and Specification in Instrumental Variable Models," *Econometric Theory*, 1993, *9* (2), 222–240.

Crump, R.K., V.J. Hotz, G.W. Imbens, and O.A. Mitnik, "Dealing with limited overlap in estimation of average treatment effects," *Biometrika*, 2009, *96* (1), 187–199.

Dalgaard, Carl-Johan, Henrik Hansen, and Finn Tarp, "On the Empirics of Foreign Aid and Growth," *Economic Journal*, 2004, *114* (496), 191–216.

Das, M., W.K. Newey, and F. Vella, "Nonparametric estimation of sample selection models," *Review of Economic Studies*, 2003, *70*, 33–58.

Dawe, D., "Can Indonesia Trust the World Rice Market?," *Bulletin of Indonesian Economic Studies*, 2008, *44*, 115–132.

Deaton, A., "Saving and Liquidity Constraints," *Econometrica*, 1991, *59*, 1221–1248.

_ , "The analysis of household surveys," 1997.

_ , *The analysis of household surveys: A microeconometric approach to development policy*, Johns Hopkins University Press, 1997.

_ **and G. Laroque**, "Competitive storage and commodity price dynamics," *Journal of Political Economy*, 1996, *104*, 896–923.

Deaton, Angus S., "Instruments, Randomization, and Learning about Development," *Journal of Economic Literature*, 2010, *48* (2), 424–455.

DeJong, David N. and Marla Ripoll, "Tariffs and Growth: An Empirical Exploration of Contingent Relationships," *Review of Economics and Statistics*, 2006, *88* (4), 625–640.

Delpierre, M., "The impact of liquidity constraints and imperfect commitment on migration decisions of offspring of rural households," *Review of Economics of the Household*, 2012, pp. 1–18.

Dickey, D. A. and W. A. Fuller, "Distribution of the estimators for autoregressive time series with a unit root," *Journal of the American Statistical Association*, 1979, *74*, 427–431.

Djankov, Simeon, Jose G. Montalvo, and Marta Reynal-Querol, "The Curse of Aid," *Journal of Economic Growth*, 2008, *13* (3), 169–235.

_ **, Rafael la Porta, Florencio López de Silanes, and Andrei Shleifer**, "Courts," *Quarterly Journal of Economics*, 2003, *118* (2), 453–517.

Doko, Firmin and Jean-Marie Dufour, "Instrument endogeneity and identification-robust tests: some analytical results," *Journal of Statistical Planning and Inference*, 2008, *138* (9), 2649–2661.

Dollar, David and Aart Kraay, "Institutions, Trade, and Growth," *Journal of Monetary Economics*, 2003, *50* (1), 133–162.

Durlauf, Steven N., Paul A. Johnson, and Jonathan R.W. Temple, "Growth econometrics," *Handbook of economic growth*, 2005, *1*, 555–677.

Easterly, William, "Can the West Save Africa?," *Journal of Economic Literature*, 2009, *47* (2), 374–447.

_ , Ross Levine, and David M. Roodman, "Aid, Policies, and Growth: Comment," *American Economic Review*, 2004, *94* (3), 774–780.

Eaton, J., S. Kortum, and S. Sotelo, "International trade: Linking micro and macro," *Advances in Economics and Econometrics: Theory and Applications, Econometric Society Monographs*, forthcoming.

Edmonds, E. and N. Pavcnik, "The Effect of Trade Liberalization on Child Labor," *Journal of International Economics*, 2003, *65*, 401–441.

Edmonds, E.V., "Child labor and schooling responses to anticipated income in South Africa," *Journal of Development Economics*, 2006, *81*, 386–414.

Eeckhout, J., "Gibrat's law for (all) cities," *American Economic Review*, 2004, *95*, 1429–1451.

Elbers, C., J.O. Lanjouw, and P. Lanjouw, "Micro–level Estimation of Poverty and Inequality," *Econometrica*, 2003, *71* (1), 355–364.

Evans, D. S. and B. Jovanovic, "An estimated model of entrepreneurial choice under liquidity constraints," *The Journal of Political Economy*, 1989, pp. 808–827.

Fernández, Carmen, Eduardo Ley, and Mark F.J. Steel, "Model uncertainty in cross-country growth regressions," *Journal of Applied Econometrics*, 2001, *16*, 563–576.

Feyrer, James and Bruce Sacerdote, "Colonialism and Modern Income," *Review of Economics and Statistics*, 2004, *91* (2), 245–262.

Field, E., R. Pande, J. Papp, and N. Rigol, "Debt Structure, Entrepreneurship, and Risk: Evidence from Microfinance," *Unpublished manuscript*, 2011.

Filmer, D. and N. Schady, "Does more cash in conditional cash transfer programs always lead to larger impacts on school attendance?," *Journal of Development Economics*, 2011, *96* (1), 150–157.

Firpo, S., "Efficient semiparametric estimation of quantile treatment effects," *Econometrica*, 2007, *75* (1), 259–276.

Fitrani, F., B. Hofman, and K. Kaiser, "Unity in Diversity? The Creation of New Local Governments in a Decentralizing Indonesia," *Bulletin of Indonesian Economic Studies*, 2005, *41*, 57–79.

Forbes, Kristin, "A Reassessment of the Relationship Between Inequality and Growth," *American Economic Review*, 2000, *90* (4), 869–887.

Frankel, Jeffrey and Andrew Rose, "An Estimate of the Effect of Common Currencies on Trade and Income," *Quarterly Journal of Economics*, 2002, *117* (2), 437–466.

_ **and David Romer**, "Does Trade Cause Growth?," *American Economic Review*, 1999, *89* (3), 379–399.

Friedman, Eric, Simon Johnson, Daniel Kaufmann, and Pablo Zoido-Lobatón, "Dodging the Grabbing Hand: The Determinants of Unofficial Activity in 69 Countries," *Journal of Public Economics*, 2000, *76* (3), 459–493.

Friedman, J. and J. A. Levinsohn, "The Distributional Impacts of Indonesia's Financial Crisis on Household Welfare: A "Rapid Response" Methodology," *National Bureau of Economic Research*, 2001.

Fuglie, K. O., "Productivity growth in Indonesian agriculture, 1961-2000," *Bulletin of Indonesian Economic Studies*, 2010b, *40*, 209–225.

Gabaix, X., "Zipf's law for cities: an explanation," *The Quarterly Journal of Economics*, 1999, *114*, 739–767.

_ , "Power Laws in Economics and Finance," *Annual Review of Economics*, 2009, *1*, 255–293.

_ **and R. Ibragimov**, "Rank-1/2: A Simple Way to Improve the Estimation of Tail Exponents," *Journal of Business and Economic Statistics*, 2011, *29*, 24–39.

Gallant, A. R. and D. W. Nychka, "Semi-nonparametric maximum likelihood estimation," *Econometrica*, 1987, *55*, 363–390.

Gallup/IOM, "Gallup World Poll: The Many Faces of Global Migration," *Special Report*, 2011.

Gaston, N. and Douglas R. Nelson, "Bridging trade theory and labour econometrics: The effects of international migration," *Journal of Economic Surveys*, 2013, *27* (1), 98–139.

Gayle, G.L. and C. Viauroux, "Root-N consistent semiparametric estimators of a dynamic panel-sample-selection model," *Journal of Econometrics*, 2007, *141*, 179–212.

Gertler, Paul, David I Levine, and Enrico Moretti, "Do microfinance programs help families insure consumption against illness?," *Health economics*, 2009, *18* (3), 257–273.

Giavazzi, Francesco and Guido Tabellini, "Economic and Political Liberalizations," *Journal of Monetary Economics*, 2005, *52* (2), 1297–1330.

Glaeser, Edward L., Rafael La Porta, Florencio López de Silanes, and Andrei Shleifer, "Do Institutions Cause Growth?," *Journal of Economic Growth*, 2004, *9* (3), 271–303.

Graham, B.S., C.C.D.X. Pinto, and D. Egel, "Inverse probability tilting for moment condition models with missing data," *The Review of Economic Studies*, 2012, *79* (3), 1053–1079.

Grogger, J. and G. Hanson, "Income maximization and the selection and sorting of international migrants," *Journal of Development Economics*, 2011, *95*, 42–57.

Guggenberger, Patrik, "The Impact of a Hausman Pretest on the Asymptotic Size of a Hypothesis Test," *Econometric Theory*, 2009, *26* (2), 369–382.

Gumbel, E. J., *Statistics of extremes*, Columbia University Press (New York), 1958.

Hahn, Jinyong and Jerry Hausman, "Estimation with valid and invalid instruments," *Annales d'Économie et de Statistique*, 2005, pp. 25–57.

_ , John Ham, and Roger Moon, "The Hausman test and weak instruments," *Journal of Econometrics*, 2011, *160* (2), 289–299.

Hall, Alastair, Atsushi Inoue, and Changmock Shin, "Entropy-Based Moment Selection in the Presence of Weak Identification," *Econometric Reviews*, 2008, *27* (5), 398–427.

Hall, Robert and Charles I. Jones, "Why Do Some Countries Produce So Much More Output Per Worker Than Others?," *Quarterly Journal of Economics*, 1999, *114* (1), 83–116.

Halliday, T., "Migration, risk, and liquidity constraints in El Salvador," *Economic Development and Cultural Change*, 2006, *54*, 893–926.

Hansen, Henrik and Finn Tarp, "Aid and Growth Regressions," *Journal of Development Economics*, 2001, *64* (2), 547–570.

Hansen, L.P., "Large sample properties of generalized method of moments estimators," *Econometrica*, 1982, pp. 1029–1054.

Hatton, T. J. and J. G. Williamson, *The age of mass migration: Causes and economic impact*, Oxford University Press, USA, 1998.

_ **and** _ , "Are Third World Emigration Forces Abating?," *World Development*, 2011, *39* (1), 20–32.

Hauk, William and Romain Wacziarg, "A Monte Carlo Study of Growth Regressions," *Journal of Economic Growth*, 2009, *14* (2), 1381–4338.

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik, "What You Export Matters," *Journal of Economic Growth*, 2007, *12* (1), 1–25.

Hayakawa, Kazuhiko, "A Simple Efficient Instrumental Variable Estimator for Panel AR(p) Models When Both N and T Are Large," *Econometric Theory*, 2009, *25* (3), 873–890.

Hayashi, F., "Tests for liquidity constraints: a critical survey and some new observations," in "Advances in Econometrics: Fifth World Congress," Vol. 2 Cambridge University Press 1987, pp. 91–120.

_ , *Econometrics*, Princeton, New Jersey: Princeton University Press, 2000.

Heckman, J. J., "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models," *NBER Chapters*, 1976, pp. 120–137.

Heckman, J.J., H. Ichimura, and P. Todd, "Matching as an econometric evaluation estimator," *Review of Economic studies*, 1998, *65* (2), 261–294.

Helpman, E., M. Melitz, and Y. Rubinstein, "Estimating Trade Flows: Trading Partners and Trading Volumes," *Quarterly Journal of Economics*, 2008, *123*, 441–487.

Holtz-Eakin, D., W. Newey, and H. S. Rosen, "Estimating Vector Autoregressions with Panel Data," *Econometrica*, 1988, *56* (6), 1371–95.

Honoré, B. E., "Trimmed LAD and least squares estimation of truncated and censored regression models with fixed effects," *Econometrica*, 1992, *60*, 533–565.

_ , “On marginal effects in semiparametric censored regression models,” *Manuscript*, 2008.

Hsieh, C-T. and P. J Klenow, “Misallocation and manufacturing TFP in China and India,” *The Quarterly Journal of Economics*, 2009, *124* (4), 1403–1448.

Hsieh, C.T., “Do consumers react to anticipated income changes? Evidence from the Alaska permanent fund,” *American Economic Review*, 2003, *93* (1), 397–405.

Hugo, G., “The Impact of the Crisis on Internal Population Movement in Indonesia,” *Bulletin of Indonesian Economic Studies*, 2000, *36*, 115–138.

_ , in T. van Naerssen, E. Spaan, and A. Zoomers, eds., *Global Migration and Development*, Psychology Press, 2008, chapter Chapter 3: International Migration in Indonesia and its Impacts on Regional Development, pp. 43–65.

Im, K.S., M.H. Pesaran, and Y. Shin, “Testing for unit roots in heterogeneous panels,” *Journal of Econometrics*, 2003, *115*, 53–74.

Imbens, Guido W., “Sensitivity to Exogeneity Assumptions in Program Evaluation,” *American Economic Review*, 2003, *93* (2), 126–132.

Imbens, G.W., “The role of the propensity score in estimating dose-response functions,” *Biometrika*, 2000, *87* (3), 706–710.

_ **and J.M. Wooldridge**, “Recent Developments in the Econometrics of Program Evaluation,” *Journal of Economic Literature*, 2009, *47* (1), 5–86.

Janvry, A. De and E. Sadoulet, “Making conditional cash transfer programs more efficient: designing for maximum effect of the conditionality,” *The World Bank Economic Review*, 2006, *20* (1), 1–29.

Jappelli, T. and L. Pistaferri, “The Consumption Response to Income Changes,” *Annual Review of Economics*, 2010, *2*, 479–506.

Jayachandran, S., “Selling labor low: Wage responses to productivity shocks in developing countries,” *Journal of Political Economy*, 2006, *114* (3), 538–575.

Johnson, D. S., J. A. Parker, and N. S. Souleles, “Household Expenditure and the Income Tax Rebates of 2001,” *The American Economic Review*, 2006, *96* (5), 1589–1610.

Jones, Ben and Benjamin Olken, “Do Leaders Matter? National Leadership and Growth Since World War II,” *Quarterly Journal of Economics*, 2005, *120* (3), 835–864.

Kaboski, J.P. and R.M. Townsend, "Policies and Impact: An Analysis of Village-Level Microfinance Institutions," *Journal of the European Economic Association*, 2005, *3* (1), 1–50.

_ **and** _ , "A Structural Evaluation of a Large-Scale Quasi-Experimental Microfinance Initiative," *Econometrica*, 2011, *79* (5), 1357–1406.

Kaplan, G. and G. L Violante, "How Much Consumption Insurance Beyond Self-Insurance?," *American Economic Journal: Macroeconomics*, 2010, *2* (4), 53–87.

Kennan, J., "Open Borders," *Review of Economic Dynamics*, forthcoming.

Kleemans, M. and J. Magruder, "Labor Market Changes in Response to Immigration: Evidence from Internal Migration Driven by Weather Shocks," *Unpublished manuscript*, 2012.

Kleiber, C. and S. Kotz, *Statistical size distributions in economics and actuarial sciences*, Wiley-Interscience, 2003.

Kleibergen, Frank, "Pivotal statistics for testing structural parameters in instrumental variables regression," *Econometrica*, 2002, *70* (5), 1781–1803.

_ **and Richard Paap**, "Generalized Reduced Rank Tests Using the Singular Value Decomposition," *Journal of Econometrics*, 2006, *133* (1), 97–126.

_ **and Sophocles Mavroeidis**, "Weak Instrument Robust Tests in GMM and the New Keynesian Phillips Curve," *Journal of Business and Economic Statistics*, 2009, *27* (2), 293–311.

Klein, P. and G. J. Ventura, "Journal of Monetary Economics," *Productivity differences and the dynamic effects of labor movements*, 2009, *56*, 1059–1073.

Klein, R.W. and R.H. Spady, "An efficient semiparametric estimator for binary response models," *Econometrica*, 1993, *61*, 387–421.

Kline, Patrick, "Oaxaca-Blinder as a reweighting estimator," *American Economic Review*, 2011, *101* (2), 532–537.

Kraay, Aart, "Instrumental Variables Regression with Honestly Uncertain Exclusion Restrictions," *World Bank Policy Research Working Paper No. 4632.*, 2008.

Kyriazidou, E., "Estimation of a panel data sample selection model," *Econometrica*, 1997, pp. 1335–1364.

Leung, S. F. and S. Yu, "On the choice between sample selection and two-part models," *Journal of econometrics*, 1996, *72* (1-2), 197–229.

Levine, D. and D. Yang, "A Note on the Impact of Local Rainfall on Rice Output in Indonesian Districts," *Unpublished manuscript*, 2006.

Levine, Ross, Norman Loayza, and Thorsten Beck, "Financial Intermediation and Growth: Causality And Causes," *Journal of Monetary Economics*, 2000, *46* (1), 31–77.

Luca, G. De and F. Peracchi, "Estimating models with unit and item nonresponse from cross-sectional sureys," *Journal of Applied Econometrics*, forthcoming.

Lundberg, Mattias and Lyn Squire, "The Simultaneous Evolution of Growth and Inequality," *Economic Journal*, 2003, *113* (487), 326–344.

Maccini, S. and D. Yang, "Under the Weather: Health, Schooling, and Economic Consequences of Early-Life Rainfall," *American Economic Review*, 2009, *99*, 1006–1026.

Mankiw, N. Gregory, "The Growth of Nations," *Brookings Papers on Economic Activity*, 1995, *1*, 275–310.

Manova, K., "Credit Constraints, Heterogeneous Firms, and International Trade," *Review of Economic Studies*, forthcoming.

Mauro, Paulo, "Corruption and Growth," *Quarterly Journal of Economics*, 1995, *110*, 681–712.

Mayda, Anna Maria, "International migration: A panel data analysis of the determinants of bilateral flows," *Journal of Population Economics*, 2010, pp. 1–26.

McCulloch, N., "Rice Prices and Poverty in Indonesia," *Bulletin of Indonesian Economic Studies*, 2008, *44*, 45–64.

McKenzie, D., "Beyond baseline and follow-up: The case for more T in experiments," *Journal of Development Economics*, 2012.

_ **and H. Rapoport**, "Network effects and dynamics of migration and inequality: Theory and evidence from Mexico," *Journal of Development Economics*, 2007, *84*, 1–24.

Melitz, M., "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 2003, *71*, 1695–1725.

Mendola, M., "Migration and Technological Change in Rural Households: Complements or Substitutes?," *Journal of Development Economics*, 2008, *85*, 150–175.

Meng, L., "Land, Household Selectivity, And Rural-Urban Migration In Hinterland China," *Unpublished Manuscript*, 2009.

_, "Can Grain Subsidy Impede Rural-Urban Migration in Hinterland China? Evidence from Field Surveys," *Unpublished Manuscript*, 2010.

MICRA, "Promoting Female Migrant Workers' Access to Finance," *Special Report of the Microfinance Innovation Center for Resources and Alternatives*, 2008.

Moreira, Marcelo J., "A conditional likelihood ratio test for structural models," *Econometrica*, 2003, *71*, 1027–1048.

Mundlak, Y., D. Larson, and R. Butzer, "Agricultural dynamics in Thailand, Indonesia and the Philippines," *Australian Journal of Agricultural and Resource Economics*, 2004, *48*, 95–126.

Munshi, K., "Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market," *Quarterly Journal of Economics*, 2003, *118*, 549–599.

Murray, Michael P., "Avoiding Invalid Instruments and Coping with Weak Instruments," *Journal of Economic Perspectives*, 2006, *20* (4), 111–132.

Naylor, R. L., W. P. Falcon, D. Rochberg, and N. Wada, "Using El Nino/Southern Oscillation climate data to predict rice production in Indonesia," *Climatic Change*, 2001, *50*, 255–265.

Newey, W. K., "Two-step series estimation of sample selection models," *Unpublished manuscript*, 1988.

_, "Two-step series estimation of sample selection models," *The Econometrics Journal*, 2009, *12*, S217–S229.

_ **and D. McFadden**, "Large sample estimation and hypothesis testing," *Handbook of Econometrics*, 1994, *4*, 2111–2245.

Newey, Whitney K. and Frank Windmeijer, "Generalized Method of Moments with Many Weak Moment Conditions," *Econometrica*, 2009, *77* (3), 687–719.

Noguer, Marta and Marc Siscart, "Trade Raises Income: A Precise and More Robust Result," *Journal of International Economics*, 2005, *65* (2), 447–460.

Okui, Ryo, "The Optimal Choice of Moments in Dynamic Panel Data Models," *Journal of Econometrics*, 2009, *151* (1), 1–16.

Olken, B.A., "Corruption and the costs of redistribution: Micro evidence from Indonesia," *Journal of Public Economics*, 2006, *90* (4), 853–870.

Orrenius, P. M. and M. Zavodny, "Self-selection Among Undocumented Immigrants from Mexico," *Journal of Development Economics*, 2005, *78*, 215–240.

Ortega, F. and G. Peri, "The causes and effects of international migrations: evidence from OECD countries 1980-2005," *NBER working paper*, 2009.

Ossa, R., "Why trade matters after all," *NBER working paper 18113*, 2012.

Papaioannou, Elias and Gregorios Siourounis, "Democratisation and Growth," *Economic Journal*, 2008, *118* (532), 1520–1551.

Parker, J. A., N. S. Souleles, D. S. Johnson, and R. McClelland, "Consumer Spending and the Economic Stimulus Payments of 2008," *American Economic Review*, forthcoming.

Persson, Torsten and Guido Tabellini, "Democracy and Economic Development: the Devil Is In the Details," *American Economic Review Papers and Proceedings*, 2006, *96*, 319–324.

_ and _ , "The Growth Effect of Democracy: Is It Heterogenous and How Can It Be Estimated?," *NBER Working Paper 13150*, 2007.

Poirier, D., "Partial Observability in Bivariate Probit Models," *Journal of Econometrics*, 1980, *12*, 209–217.

Qian, Nancy, "Missing Women and the Price of Tea in China: The Effect of Sex-Specific Earnings on Sex Imbalance," *Quarterly Journal of Economics*, 2008, *123*, 1251–1285.

Rajan, Raghuram and Arvind Subramanian, "Aid and Growth: What Does the Cross-country Evidence Really Show?," *Review of Economics and Statistics*, 2008, *90* (4), 643–665.

Rochina-Barrachina, M. E., "A new estimator for panel data sample selection models," *Annales d'Économie et de Statistique*, 1999, *55-56*, 153–181.

Rodríguez, Francisco and Dani Rodrik, "Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence," *NBER Macroeconomics Annual*, 2001, *15*, 261–325.

Rodrik, Dani and Romain Wacziarg, "Do Democratic Transitions Produce Bad Economics Outcomes?," *American Economic Review Papers and Proceedings*, 2005, *95*, 50–56.

Roodman, David, "How to Do xtabond2: An Introduction To Difference and System GMM in Stata," *Stata Journal*, 2009, *9* (1), 86–136.

— , "A Note on the Theme of Too Many Instruments," *Oxford Bulletin of Economics and Statistics*, 2009, *71* (1), 135–158.

Rose, Andrew K., "Size Really Doesn't Matter: In Search of a National Scale Effect," *Journal of the Japanese and International Economies*, 2006, *20* (4), 482–507.

Rosenzweig, M. and K. I. Wolpin, "Natural natural experiments," *Journal of Economic Literature*, 2000, *38* (4), 827–874.

Sala–i–Martin, Xavier, "I Just Ran Two Million Regressions," *American Economic Review*, 1997, *87* (2), 178–183.

— , Gernot Doppelhofer, and Ronald I. Doppelhofer, "Determinants of Long-term Growth: A Bayesian Averaging of Classical Estimates (BACE) Approach," *American Economic Review*, 2004, *94* (4), 813–835.

Simpson, N. B. and C. Sparber, "The Short-and Long-Run Determinants of Unskilled Immigration into US States," *Unpublished Manuscript*, 2010.

Skoufias, E., A. Narayan, K. Kaiser, and B. Dasgupta, "Electoral Accountability, Fiscal Decentralization and Service Delivery in Indonesia," *World Bank Policy Research Working Paper*, 2010.

Solon, G., S. J. Haider, and J. Wooldridge, "What Are We Weighting For?," *NBER Working Paper*, 2013.

Soo, K. T., "Zipf's law for cities: a cross-country investigation," *Regional Science and Urban Economics*, 2005, *35*, 239–263.

Spaan, E., "Taikongs and Calos: The Role of Middlemen and Brokers in Javanese International Migration," *International Migration Review*, 1994, *23*, 93–113.

Spolaore, Enrico and Romain Wacziarg, "Borders and Growth," *Journal of Economic Growth*, 2005, *10* (4), 331–386.

Stock, James H. and Jonathan H. Wright, "GMM with Weak Identification," *Econometrica*, 2000, *68* (5), 1055–1096.

— and Motohiro Yogo, "Testing for Weak Instruments in Linear IV Regression," in James H. Stock and Donald W. K. Andrews, eds., *Identification and Inference for Econometric Models: Essays in Honor of Thomas J Rothenberg*, New York: Cambridge University Press, 2005.

Suryahadi, A., W. Widyanti, R. Prama Artha, D. Perwira, and S. Sumarto, "Developing a Poverty Poverty Map for Indonesia: A Tool for Better Targeting in Poverty Reduction and Social Protection Programs, Book 1: Technical Report," *SMERU Research Report*, 2005.

Tavares, Jose and Romain Wacziarg, "How Democracy Affects Growth," *European Economic Review*, 2001, *45* (3), 1341–1375.

Temple, Jonathan, "The New Growth Evidence," *Journal of Economic Literature*, 1999, *37* (1), 112–156.

Timmer, C. P., "Causes of High Food Prices," *Asian Development Bank Working Paper*, 2008, *128.*

VanWey, L. K., "Land Ownership as a Determinant of International and Internal Migration in Mexico and Internal Migration in Thailand," *International Migration Review*, 2005, *39*, 141–172.

Voitchovsky, Sarah, "Does the Profile of Income Inequality Matter for Economic Growth?," *Journal of Economic Growth*, 2005, *10* (3), 273–296.

Vothknecht, M. and S. Sumarto, "Violent Conflict and Economic Growth: Evidence from Indonesian Villages," *Unpublished manuscript*, 2009.

Warr, P., "Food policy and poverty in Indonesia: a general equilibrium analysis," *Australian Journal of Agricultural and Resource Economics*, 2005, *49* (4), 429–451.

_ , "The transmission of import prices to domestic prices: an application to Indonesia," *Applied Economics Letters*, 2008, *15*, 499–503.

Wilde, J., "Identification of multiple equation probit models with endogenous dummy regressors," *Economics Letters*, 2000, *69*, 309–312.

Wimanda, R. E., "Price variability and price convergence: Evidence from Indonesia," *Journal of Asian Economics*, 2009, *20*, 427–442.

Windmeijer, Frank, "A finite sample correction for the variance of linear efficient two-step GMM estimators," *Journal of Econometrics*, 2005, *126*, 25–51.

Wolpin, K. I., "A new test of the permanent income hypothesis: the impact of weather on the income and consumption of farm households in India," *International Economic Review*, 1982, *23* (3), 583–594.

Wooldridge, J. M., "Selection corrections for panel data models under conditional mean independence assumptions," *Journal of Econometrics*, 1995, *68* (1), 115–132.

World Bank, "Improving Access to Financial Services in Indonesia," *World Bank Working Paper*, 2009.

_ , "Enhancing Access to Finance for Indonesian Overseas Migrant Workers: Evidence from a Survey of Three Provinces," *World Bank Working Paper*, 2010.

Yamagata, T., "The small sample performance of the Wald test in the sample selection model under the multicollinearity problem," *Economics Letters*, 2006, *93* (1), 75–81.

Yang, D., "Why Do Migrants Return to Poor Countries? Evidence from Philippines Migrants' Exchange Rate Shocks," *Review of Economics and Statistics*, 2006, *88*, 715–735.

_ , "International Migration, Remittances, and Household Investments: Evidence from Philippine Migrants' Exchange Rate Shocks," *The Economic Journal*, 2008a, *118*, 591–630.

_ , "Risk, Migration, and Rural Financial Markets: Evidence from Earthquakes in El Salvador," *Social Research*, 2008b, *75*, 955–992.

_ , "Migrant Remittances," *Journal of Economic Perspectives*, 2011, *25* (3), 129–52.

_ **and H. Choi**, "Are remittances insurance? Evidence from rainfall shocks in the Philippines," *The World Bank Economic Review*, 2007, *21* (2), 219–248.

Yen, S. T., "A multivariate sample-selection model: estimating cigarette and alcohol demands with zero observations," *American Journal of Agricultural Economics*, 2005, *87* (2), 453–466.

Yogo, Motohiro, "Estimating the elasticity of intertemporal substitution when instruments are weak," *Review of Economics and Statistics*, 2004, *86*, 797–810.

Zellner, A. and T. H. Lee, "Joint estimation of relationships involving discrete random variables," *Econometrica: Journal of the Econometric Society*, 1965, pp. 382–394.

Zivot, E. and D. W. K. Andrews, "Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis," *Journal of Business and Economic Statistics*, 2002, *20*, 25–44.