Title
Essays in Applied Microeconomics

Author
McCully, Brett Alexander

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Essays in Applied Microeconomics

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Economics

by

Brett Alexander McCully

2021

ABSTRACT OF THE DISSERTATION

Essays in Applied Microconomics

by

Brett Alexander McCully

Doctor of Philosophy in Economics

University of California, Los Angeles, 2021

Professor Jonathan E. Vogel, Chair

This dissertation includes three essays in applied microeconomics. In Chapter 1, I detail how immigrants facilitate illegal drug trafficking in the context of Spain. I do so by drawing upon novel data on international illegal drug confiscations and by implementing a gravity equation estimation strategy. I find that immigrants without legal status primarily drive immigrant-induced trafficking of illegal drugs, and that immigrant social connections back to their home country are also relevant for explaining illegal trafficking. In Chapter 2, with coauthors I find that consumers rarely use funds from equity extraction to purchase a car directly, even during the mid-2000s' housing boom; this finding holds across three nationally representative household surveys. We find in credit bureau data that equity extraction does lead to a statistically significant increase in auto loan originations, consistent with equity extraction easing borrowing constraints in the auto loan market. This channel, though, accounts for only a tiny share of overall car purchases.

In Chapter 3, with coauthors I explore the long-run effects of the quasi-randomized Matlab Maternal and Child Health and Family Planning (MCH-FP) program introduced in a rural area of Bangladesh from 1977-1988 on firm and farm profitability 35 years later. Using a rich data set including two followup surveys with unusually low rates of attrition, we estimate that the MCH-FP raised firm and farm productivity in the long-run.

The dissertation of Brett Alexander McCully is approved.

Felipe M. Goncalves

Emily Karen Weisburst

Randall S. Kuhn

Adriana Lleras-Muney

Pablo David Fajgelbaum

Jonathan E. Vogel, Committee Chair

University of California, Los Angeles

2021

I dedicate this dissertation to Katie, Cookie, and my family.

Contents

List of Figures

List of Tables

ACKNOWLEDGMENTS

I am deeply grateful for the guidance and mentorship given by my committee, without which this dissertation would not be possible. During my PhD I received generous funding from UCLA's Graduate Division, an NICHD Traineeship, and the Institute on Global Conflict and Cooperation's Dissertation Fellowship.

Chapter 2 is the accepted version of, "How Much Are Car Purchases Driven by Home Equity Withdrawal? Evidence from Household Surveys" (coauthored with Daniel J. Vine and Karen M. Pence), 2019, Vol. 51, No. 5, *Journal of Money, Credit, and Banking*, with the published version available at https://dx.doi.org/10.1111%2Fjmcb.12595. Karen Pence conceived of the research idea and empirical design and contributed to writing the paper and Daniel Vine cleaned and analyzed data and contributed to writing the paper.

Chapter 3 is being prepared for submission for publication, and is coauthored with Tania Barham (conceived of research idea, helped design the survey), Randall Kuhn (conceived of research idea, helped design the survey, and guided the empirical analysis), and Patrick Turner (processed and cleaned data).

VITA

Brett Alexander McCully

EDUCATION

Master of Arts (2016) in Economics, University of California in Los Angeles, Los Angeles, California.

Bachelor of Arts in Mathematics and Economics (2013), University of Virginia, Charlottesville, Virginia.

PUBLISHED WORK

"How Much Are Car Purchases Driven by Home Equity Withdrawal? Evidence from Household Surveys" (with Daniel Vine and Karen Pence), 2019, Vol. 51, No. 5, *Journal of Money, Credit, and Banking.*

"Is Underemployment Underestimated? Evidence from Panel Data" (with Geng Li), May 2016, FEDS Notes.

ACADEMIC POSITIONS

Resident Affiliate, California Center for Population Research, University of California in Los Angeles, 2016–present.

Teaching Fellow, Department of Economics, University of California in Los Angeles, 2016– 2017.

NIH-NICHD Trainee, California Center for Population Research, University of California, Los Angeles, 2017-2018 and 2019-2020.

Research Assistant to Professor Randall Kuhn, University of California, Los Angeles, 2019.

Research Assistant to the California Policy Lab, 2017.

HONORS, SCHOLARSHIPS, AND FELLOWSHIPS

Dissertation Year Fellowship, Graduate Division, University of California in Los Angeles, 2020-2021.

Institute on Global Conflict and Cooperation Dissertation Fellowship, 2020-2021.

European Studies Fellowship, UCLA, 2020-2021.

NIH-NICHD Predoctoral Fellowship, University of California in Los Angeles, 2017-2018 and 2019-2020.

Best Paper Award, International Economics Proseminar, 2019-2020.

Graduate Summer Research Mentorship, UCLA, 2017.

Pauley Fellowship, UCLA, 2015-2016 and 2018-2019.

Chapter 1

Immigrants, Legal Status, and Illegal Trade

Brett A. McCully[1]

Abstract

Nearly $2 trillion worth of illegal goods are trafficked across international borders every year, generating violence and other social costs along the way. Some have controversially linked illegal trafficking to immigrants, yet an appropriate immigration policy response is unclear. In this paper, I use novel data on nearly 10,000 confiscations of illegal drugs in Spain to study how immigrants and immigration policy affect the pattern and scale of illegal drug trafficking. To identify the causal effect of immigrants on trafficking, I construct an instrumental variable that interacts variation in total immigrant inflows into Spain by origin country with the fraction of immigrants inflowing into a province. I find that a 10% increase in the population of immigrants from a given origin country relative to the mean raises the value of drugs trafficked from the origin country confiscated in a given province by 12%. Moreover, this relationship is driven entirely by immigrants without legal status. To better understand the role of legal status, I exploit an extraordinary regularization of nearly half a million immigrants in 2005. Event study estimates suggest that granting immigrants legal status results in a long-run decline in drug trafficking, corresponding to the acquisition of citizenship by the immigrants.

[1]UCLA, Department of Economics. Click here for the most recent version of this paper. I am especially grateful to Jonathan Vogel, Pablo Fajgelbaum, Felipe Goncalves, Randall Kuhn, Adriana Lleras-Muney, and Emily Weisburst for advice and encouragement. I thank Wookun Kim for helpful conversations, and seminar participants at UCLA, Collegio Carlo Alberto, the CLEAN Unit at Bocconi, the OCC, CFPB, GFLEC, EGSC, EWMES, APPAM, PAA, and DemSemX for helpful comments. I owe special thanks to Ariadna Jou for assistance in contacting the Spanish government. I acknowledge financial support from CCPR's Population Research Infrastructure Grant P2C from NICHD: P2C-HD041022, CCPR's Population Research Training Grants T32 from NICHD: T32-HD007545, and the IGCC. This work used computational and storage services associated with the Hoffman2 Shared Cluster provided by UCLA Institute for Digital Research and Education's Research Technology Group. All errors are my own.

1.1 Introduction

Many illegal goods are not produced where they are consumed, resulting in the trafficking of nearly \$2 trillion of illegal goods across international borders annually—worth 10% of the value of legal global merchandise trade (Mavrellis, 2017). Violence often follows in the wake of illegal trafficking, and further costs to society occur when the illegally trafficked goods—particularly illegal drugs—are consumed (NDIC, 2011). This illegal trafficking often relies on informal connections and social ties to facilitate the movement of goods without binding contracts (Marsh et al., 2012).

One controversial but untested opinion holds that immigrants, particularly those without legal status, facilitate the trafficking of illegal goods from their origin country to their host region.[2] Immigrants' social connections to their origin country may make arranging for imports and exports (legal or illegal) easier (Rauch and Trindade, 2002; Combes et al., 2005; Dunlevy, 2006). In addition, immigrants without legal status are prevented from working in the formal sector, thereby reducing their earnings relative to their legal counterparts (Kossoudji and Cobb-Clark, 2002; Kaushal, 2006; Simón et al., 2014; Sanromá et al., 2015). The Becker-Ehrlich model of crime (Becker, 1968; Ehrlich, 1973) suggests that this differential in earnings will result in a higher propensity to participate in financially motivated illegal activities, such as trafficking illegal goods.

In this paper, I estimate how immigrants and immigration policy affect the trafficking of one of the most consequential illegal goods: illegal drugs. I use novel data on drug confiscations from Spain and exogenous variation in immigrant populations to show that immigrants without legal status have a large positive causal effect on the trafficking of illegal drugs from the immigrants' countries of origin. I find no effect of legal immigration on illegal drug trafficking. Because there may be characteristics of immigrants that shape selection into legal status and into drug trafficking, I estimate the dynamic effects of a mass immigrant regularization policy. I find that granting immigrants legal status results in a long-run decline in drug trafficking, corresponding to immigrants acquiring citizenship.

The main contribution of this paper is to provide the first causally identified estimates of the effect of immigrants on illegal trafficking and the first exploration of mechanisms that generate this relationship. Credibly establishing a causal relationship between im-

[2]Several notable politicians have made this claim. Donald Trump suggested in 2015 that Mexican immigrants were "bringing drugs [and] crime" into the United States. Then-presidential candidate Sebastian Piñera in 2017 blamed Chile's immigration laws for "importing problems like delinquency, drug trafficking and organized crime" (Esposito and Iturrieta, 2017). In addition, the European Union High Representative for Common Foreign and Security Policy argued in 2003 that, "massive flow[s] of drugs and migrants are coming to Europe and [will] affect its security. These threats are significant by themselves, but it is their combination that constitutes a radical challenge to our security" (Solana, 2003). More broadly, in both the United States and European rounds of the Transatlantic Trends survey, respondents blame irregular immigrants for increasing crime much more than they blame regular immigrants.

migrants without legal status and drug trafficking is challenging for two reasons. First, the illegal nature of trafficking and undocumented immigration makes measurement of these two phenomenons difficult. Second, other factors (such as geography) may affect both the distribution of immigrant populations and illegal drug trafficking.

To make progress on the difficulty in measuring illegal drug trafficking, I use detailed data on drug confiscations that include information on which country the drugs were trafficked from. In particular, I use a database of individual drug confiscations as a proxy for actual drug flows in the context of Spain, a country with high-quality reporting of data on drug confiscations. These data report where the drug confiscation occurred within Spain, from which country the drugs were trafficked, and, if available, to which country the drugs were intended to be trafficked, thus providing insight into the region-to-region flows of illegal drugs. To validate that this indirect measure captures variation in actual flows of illegal goods, I compare confiscations to survey-based measures of drug use and availability at the province level. I find that more confiscations correspond to more drug use and availability.

Spain provides a unique context to study whether and how immigrants and immigrant legal status affect the flow of illegal drugs. In particular, Spain is a major hub for cocaine and cannabis trafficking into Europe. The country has also experienced substantial immigration in recent decades, much of it irregular.

I exploit unique institutional features in Spain that facilitate the measurement of irregular immigrant populations. Unlike the United States and other European countries, immigrants to Spain can obtain healthcare and other government benefits regardless of their legal status in exchange for registering with the local population registry. Comparing local population registries with counts of permits for legal residency leads to a straightforward estimation of the size of the irregular immigrant population (González-Enríquez, 2009; Gálvez Iniesta, 2020).

To make progress on causal identification, I estimate a gravity equation, the workhorse model in the international trade literature used to explain the volume of trade flowing from one region to another (Tinbergen, 1962; Head and Mayer, 2014). In particular, I estimate a gravity equation of illegal drug flows from a given origin country on the number of immigrants from the country living in a given Spanish province. Because I observe origins and destinations of both drugs and immigrants, I can flexibly control for observed and unobservable features of each country and each Spanish province using country and province fixed effects. These fixed effects absorb variation in either law enforcement activity directed towards specific nationalities in Spain (in the case of the country fixed effect) or variation in law enforcement efficacy in confiscating drugs across provinces (in the case of the province fixed effect).

There may still be factors at the country-province pair level that drive both drug trafficking and immigration from the country to the province. For example, Morrocan immigrants and Moroccan drug traffickers may be drawn to Barcelona for its familiar Mediterranean climate. To address this potential endogeneity, I adapt I adapt the instrumental variables approach developed by Burchardi et al. (2019) to generate exogenous variation in the number of immigrants from a given country living in a given Spanish province. The instrument relies on the intuition that immigrants from origin country o are likely to settle in Spanish province d if many immigrants from o are arriving in Spain at the same time that many immigrants are settling in d. In particular, the instrument interacts the "pull" of Spanish province d to immigrants—measured as the share of immigrants in a given decade settling in d—with the "push" to immigrate from origin country o—measured as the number of immigrants from o entering Spain in a given decade.

I find that a higher immigrant population from a given origin country facilitates the import and re-export of illegal drugs from that origin country. For an average Spanish province, I find that a 10% increase in the number of immigrants relative to the mean from a given origin country raises the likelihood that illegal drugs trafficked from the origin country will be confiscated locally by 0.5 percentage points, and raises the market value of confiscated drugs coming from the origin country by 12%. Similarly, a 10% increase in the number of immigrants relative to the mean from a given origin country raises the likelihood that drugs intended for re-export to the immigrants' home country will be confiscated locally by 0.4 percentage points and raises the value of drugs intended for re-export to the immigrants' home country by 7%.

These main results are robust to a range of alternative specifications and sampling choices. I relax the functional form assumption in my baseline specification, separately using non-linear generalized method of moments and non-parametric estimation methods, and find results consistent with my baseline estimation. In addition, no single drug or region drives my baseline result, as I find consistent effects when leaving out individual origins, destinations, and drugs. To gauge the reasonableness of the estimated magnitude, I compare coefficients between a gravity model of illegal trafficking and a gravity model of legal trade and find similar effect sizes.

I argue that immigrants' social connections to their origin country primarily drives the bilateral immigrant-trafficking relationship that I estimate. This is consistent with the qualitative evidence that immigrants reduce information frictions and transaction costs for imports and exports. In addition, I find that immigrants raise re-exports of drugs, a margin where immigrants' demand for drugs should not drive the results. An alternative explanation is that immigrants may prefer to consume goods from their home country (Bronnenberg et al., 2012; Atkin, 2013). However, product differentiation of illegal drugs across trafficking (not production) origins is unlikely to occur in the context of drug markets. In addition, I find that immigrants consume drugs at significantly lower rates than native-born Spaniards, and immigrants raise re-exports at similar magnitudes as they raise imports.

A competing explanation for my baseline results is that the intensity with which law enforcement conducts drug enforcement activities is affected by the size of the local immigrant population. Due to the origin country and province fixed effects in my baseline specification, this competing explanation must operate at the origin country-by-Spanish province level. I take two approaches to rule out that such enforcement intensity variation drives my baseline results. First, I combine my baseline estimates of the effect of immigrants on drug confiscations with a back-of-the-envelope estimate of the fraction of illegal drugs coming into Spain which are confiscated by the authorities. I find that an implausibly large responsiveness of enforcement intensity to immigration is required to explain my quantitatively large baseline estimates. Second, while in my baseline estimation I assume that enforcement intensity does not co-vary with the immigrant population. I test this assumption by focusing on the extensive margin of drug trafficking. I still find a large positive effect of immigrants on confiscations at the extensive margin of trafficking, suggesting that enforcement intensity cannot fully explain my baseline results.

I also find that general equilibrium responses, including changes in the participation of the native-born in drug markets, do not completely offset the effect of immigrants on trafficking. I assess the strength of these general equilibrium responses by estimating the effect of immigrants on additional measures of drug market activity at the province level. I find that an increase in the immigrant population in a province (across all origin countries) raises the value of drugs confiscated locally.

I estimate the effect of immigrants on drug trafficking separately by immigrant legal status using the gravity specification. I find that my baseline estimates are driven entirely by irregular immigrants. To achieve causal identification, I interact the leave-out push-pull instrument from the baseline estimation with a predicted propensity for immigrant irregularity at the origin country-province level. I predict irregularity for a country-province pair in 2011 using the share of immigrants from the country and outside the region of the province back in 2003.

Unobserved immigrant characteristics, such as a propensity for illegal behavior, may drive immigrants into both irregular status and drug trafficking. These differences in the composition immigrants by legal status at the origin country-province level may partly explain my instrumented gravity estimates. To better understand the effects of legal status on trafficking, I exploit a major immigrant regularization program implemented in 2005. This program resulted in nearly half a million immigrants receiving legal status and also put regularized immigrants on the path to citizenship. Immigrants are eligible for citizenship after living in Spain continuously and legally for a number of years depending on their country of origin.

I find that the 2005 mass immigrant regularization program reduced drug trafficking significantly shortly after immigrants became eligible to become Spanish citizens, but not before. The lack of an immediate effect of regularization on drug trafficking is consistent with Pinotti (2017), as the program-eligible immigrants had pre-existing attachments to the formal labor market.A back-of-the-envelope calculation leveraging the gravity estimates suggests that an alternative policy in which regularization was not conditional on pre-existing attachment to the formal labor market would have reduced drug confiscations by as much as 20 percent.

This paper provides the first causally identified estimates of the effect of immigrants and immigrant legal status on illegal trafficking. Related work by Berlusconi et al. (2017), Giommoni et al. (2017), and Aziani et al. (2019) uses country-pair level data on drug confiscations to assess how immigrant population at the country-pair level correlates with drug confiscations. I make several advancements relative to this literature. First, I use credibly exogenous variation in bilateral immigrant population. Second, I include origin and destination fixed effects to control for observed and unobserved factors at the region-level that shape immigration and trafficking. Third, I exploit within-country variation, which allows me to control country-pair level factors. Finally, I explore the underlying mechanisms that drive the observed immigrant-trafficking relationship and the resulting immigration policy implications.

This article contributes to the debate on the costs and benefits of immigration and on which immigration policies host countries should implement. Much of the literature on the consequences of immigration has focused on labor market outcomes.[3] A separate literature has estimated the effect of immigrants on legal trade (Gould, 1994; Head and Ries, 1998; Rauch and Trindade, 2002; Combes et al., 2005; Cohen et al., 2017; Parsons and Vézina, 2018). This paper expands upon this literature by looking at a new outcome—illegal trade—changing as a result of immigration and by showing that the legal status regime of the host country is crucial for shaping this relationship.

My work complements existing studies on the effect of immigrants on crime. I provide evidence for a new mechanism linking immigration and crime: immigrants' social connections to their home country. Prior research on immigration and crime tends to focus on the labor market opportunities available to immigrants (Bell et al., 2013; Spenkuch, 2014; Pinotti, 2017; Freedman et al., 2018). I also show the potential for long-run effects of immigrant legalization, in part due to immigrant naturalization, whereas prior work focuses on short-run effects.

I also expand upon the literature on the economics of illegal trade by studying the trafficking of illegal drugs, one of the most consequential of illegally smuggled goods.[4] I

[3] See, for example, Card (2001), Friedberg (2001), Borjas (2003), Dustmann et al. (2013), and Monras (2020). For a recent review of the literature, see Dustmann et al. (2016).

[4] A key distinction between past studies on the economics of drug trafficking and the present paper

follow a strand of mostly theoretical papers on the economics of smuggling (Bhagwati and Hansen, 1973; Grossman and Shapiro, 1988; Thursby et al., 1991). In more recent work, Fisman and Wei (2009) empirically study the smuggling and mis-invoicing of cultural goods, and Akee et al. (2014) estimates the determinants of human trafficking.

This paper proceeds as follows. Section 1.2 introduces the data and validates the drug confiscations data as a proxy for actual drug flows. Section 1.3 presents my empirical strategy and results. Section 1.4 discusses enforcement intensity and general equilibrium responses, and Section 1.5 discusses the role for immigration policy. Section 1.6 concludes.

1.2 Background and Measurement of Drug Trafficking

1.2.1 Background

Illegal Drugs. The most commonly consumed illegal drugs around the world are cannabis, opioids, amphetamines and prescription stimulants, ecstasy, and cocaine, ranked by number of users in 2018 (p.7, UNODC, 2020b). Cannabis and cocaine are the primary drugs trafficked in Spain. The country serves as an key entry point to Europe for these drugs.[5]

Illegal drugs typically pass through many countries between their production location and final consumption location. Cocaine, for example, is grown exclusively in three countries in the world: Colombia, Peru, and Bolivia. While the United States and Europe represent the primary consumption regions in the world, cocaine passes through intermediary countries such as Mexico or West Africa on the way to these markets.[6]

Cannabis, by contrast, "is produced in almost all countries worldwide."[7] Nevertheless, a large amount of cannabis is still trafficked across international borders, although it tends to remain in the same region.[8]

In Spain, confiscations of domestic cannabis plants (Alvarez et al., 2016) are quite small compared to the amount of cannabis confiscated arriving from abroad. Amphetamines can also be produced locally, but are a small part of the market, with only 2% of drug treatment patients seeking help for an amphetamine addiction. This fraction is roughly in line with the share of amphetamines in total confiscations.[9]

Due to the intermediary-intensive nature of trafficking, social connections between countries may facilitate trafficking routes. For example, in a set of interviews in the

is that I look at *bilateral*, rather than region-specific, determinants of drug trafficking. Other studies have looked at the consequences of law enforcement crackdowns on drug cultivation (Abadie et al., 2014; Mejía et al., 2017) and violence (Castillo et al., 2020). A notable exception is Dell (2015), who estimates how crackdowns shape violence and drug trafficking networks. However, Dell (2015) lacks data on the bilateral flows of illegal drugs.

[5] See https://www.emcdda.europa.eu/countries/drug-reports/2019/spain/drug-markets_en.

[6] UNODC (p. 30, 2020a).

[7] UNODC (p. 67, 2020a).

[8] UNODC (p. 71-73, 2020a).

[9] See https://www.emcdda.europa.eu/countries/drug-reports/2019/spain_en.

United Kingdom conducted by Matrix Knowledge Group (2007), jailed traffickers shared the importance of social ties. Most recruiting of workers in the drug trafficking business occurred within one's social network[10], and traffickers also noted examples in which a shared nationality raised trust between individuals seeking to conduct illegal trade transactions.[11]Proximity to immigrants from a variety of drug source countries was seen as advantageous as it reduced search costs.[12] In the context of legal trade, Rauch and Trindade (2002) note that punishment of cheating firms within a migrant network can facilitate trade given incomplete contracts, which bear particular relevance for the case of illegal transactions.

Immigration. Spain has experienced tremendous immigration in recent decades. Between 1991 and 2011, the share of immigrants in Spain's population rose from below 1% to well over 10% as shown in Figure 1.10, representing "the highest rate of growth of the foreign-born population over a short period observed in any OECD country since the Second World War" (OECD, 2010).

Immigrants without legal status, or irregular immigrants, are a common feature of immigration in Spain. Irregular immigrants are defined as those living in the country without a residency permit, and they generally enter Spain through legal means (González-Enríquez, 2009). These include immigrants who overstay their tourist visas and stay in Spain beyond the terms of their temporary residence permits.[13] Moreover, irregular immigration is a common phenomenon in Spain among immigrants. Surveys of immigrants in Spain have found that nearly 50% of immigrants are irregular (Pajares, 2004; Yruela and Rinken 2005). Díez Nicolás and Ramírez Lafita (2001) found that 83% of immigrants had arrived in Spain without a work permit but nevertheless began to work or look for a job.

Concurrent with its high levels of immigrant irregularity has been Spain's relatively more generous provision of public services to irregular immigrants as well as providing a path to regular status and thereafter to citizenship. For example, the country regularly provided legal status to hundreds of thousands of irregular immigrants in waves of regularizations between 2000 and 2005. In addition, irregular immigrants are eligible for

[10]"A number of interviewees indicated that the importance of trust meant that they only recruited employees [for their smuggling organization] largely through their existing social networks." (Marsh et al., 2012)

[11]For example, "L-15 [a convicted drug trafficker] was from Ghana. In 2000 he was approached by a Ghanian friend to manage his drug business in the United Kingdom. He was trusted by the dealers he had to manage because they knew his family in Ghana." (Marsh et al., 2012)

[12]For example, one convicted trafficker said that to import cocaine into the United Kingdom, "You need to know someone in the West Indies but this is not difficult to do. London is multicultural, you can meet a contact." Matrix Knowledge Group (2007)

[13]Irregular immigrants who enter Spain via either crossing the Strait of Gibralter by boat or by illegally entering the Spanish North African cities of Ceuta or Mellila are a small fraction of irregular immigrants, though they garner a disproportionate share of press coverage (González-Enríquez, 2009).

access to the country's public healthcare and education systems so long as they register with the local population registry. These benefits create a strong incentive for irregular immigrants to register, a fact that I leverage to measure irregular migration prevalence in Section 1.5.1.[14]

Obtaining legal status puts immigrants on the path to citizenship. Immigrants must live in Spain continuously and legally for ten years before they can apply for naturalization. For immigrants from Latin America, this requirement drops to two years. In addition, immigrants must meet various assimilation and "good citizen" requirements, such as Spanish language fluency and not comitting crimes.

1.2.2 Drug Trafficking Data Description

Data limitations typically complicate the study of illegal activity. In the context of drug trafficking, I use data on confiscations of illegal drugs by law enforcement to proxy for actual illegal drug flows. To validate that drug confiscations capture variation in actual flows of illegal goods, I compare confiscations to survey-based measures of drug availability and use them at the province level.

I use a database of individual drug confiscation events to proxy for actual drug flows in the context of Spain, a country with high-quality reporting of drug confiscations. Using enforcement-based measures as a proxy for illegal and therefore hard-to-observe activity is typical in the study of crime. For example, Dell (2015) uses confiscations of illegal drugs in a region as a proxy for the amount of illegal drugs flowing through the region.[15] Similarly, Dube et al. (2016) uses the number of opium poppy and cannabis plants eradicated as a proxy for cultivation.

I measure drug confiscations using a novel dataset of individual wholesale-level confiscations events compiled by the United Nations Office of Drugs and Crime (UNODC). An observation in these data is a single drug confiscation event and details the drug type, the amount confiscated, the country from which the drugs were trafficked, and the location of the confiscation. By including both the locality of a confiscation and its country of departure, I observe the bilateral linkage for each confiscation event. A subset of confiscations lists the intended destination country of the confiscated drugs. To transform quantities confiscated in dollar amounts, I use illegal drug prices reported by the Centre

[14] The population registry is an imperfect measure for several reasons. First, municipalities differ in their documentation requirements for registration and the degree to which they notify immigrants that they must re-register every two years. In addition, according to González-Enríquez (2009), sex workers and immigrants from China are less likely to register due to deportation fears. This will impact my estimation strategy only if there is a bilateral-specific measurement error term, so origin country-specific immigrant behaviors common across all provinces, or destination province policies common across all origins will be controlled for by the origin and destination fixed effects.

[15] Whereas my data on drug confiscations are at the bilateral (region-to-region) level, Dell (2015) uses confiscations aggregated to the region-level.

of Intelligence against Organized Crime at the Spanish Ministry of the Interior.[16]

I primarily use confiscations reported by Spain due to their high quality.[17] These data are compiled in Spain's Statistical System of Analysis and Evaluation on Organized Crime and Drugs, a centralized repository of information on organized crime and the illegal drug trade. This database is filled out by three national law enforcement agencies: the National Police, the Guardia Civil, and the Customs and Excise Department. These agencies report both confiscations made by their own personnel as well as by those conducted in concert with, or exclusively by, local law enforcement authorities.

Country of origin and intended destination for each drug confiscation in the dataset is assigned based on subsequent investigation, where country of origin refers to the most recent foreign country the drugs had been in (not necessarily the country in which they were produced). For some drug interdictions, assignment of origin and destination country is fairly straightforward. For drugs confiscated from airline passengers upon arrival at an airport, the origin country is the passenger's departure country and destination country is the passenger's ultimate destination on their travel itinerary. For drugs confiscated from cargo ship containers, a range of documents are checked for country of origin and intended destination, including the bill of lading, the commercial invoice, the certificate of origin, customs clearance forms, and the relevant letter of credit. In the case of "narco-boats" that transport hashish resin in the Strait of Gibraltar, their country of origin is considered to be Morocco unless proven otherwise.

For less straightforward cases, such as the case of drug gangs transporting cocaine intercepted in the Atlantic Ocean off the Galician coast, the country of origin and destination is determined based on additional information such as suspect and witness interviews and coordination with law enforcement agencies in the suspected origin and destination countries. If a person is arrested within Spain for drug trafficking but is outside an airport or port, the country of origin of the drugs will be determined on the basis of the investigation carried out, including any statements made by the arrested person. [18]

Four facts emerge when looking at the data on confiscations in Spain. First, nearly all drugs confiscated by Spanish authorities are cocaine or cannabis, with negligible amounts of amphetamines and heroin as shown in Figure 1.11. Second, the distribution of drug

[16] Specifically, these are prices in dollars for 2012 for heroin, cocaine, amphetamines, and cannabis as reported by Spain to the UNODC. I assume prices are uniform across origins and destinations.

[17] Reporting drug confiscations to the UNODC is voluntary. I focus on Spain, a country that reports a large number of drug confiscations to the UNODC annually (see Figure 1.9) and reports substantially higher quality data than other countries. For example, Spain reports at high rates fields typically missing from reports by other countries, such as the hiding place of confiscated drugs, the installation where law enforcement found the drugs, the mode of transport, and the routing of the drugs. Between 2011 and 2016, confiscation events from Spain were missing these fields for only 20% of events, while the fraction of these variables missing rose to 33% when turning to other countries. In the same time period, Spain reported the highest number of confiscations of any country.

[18] The preceding description is based on discussions with representatives from the Spanish Ministry of the Interior.

confiscation amounts is right skewed as shown in Figure 1.12, with many moderate-sized confiscations (the median confiscation value is $43,796) and a few huge confiscations (the mean confiscation value is $593,795). Third, Spain imports cannabis almost exclusively from Morocco and cocaine from Latin America, as shown in Figure 1.13, and Spain re-exports drugs primarily to the rest of Europe and the Mediterranean region. Finally, there is substantial spatial variation across Spain of the import and export of illegal drugs, as shown in Figures 1.15 and 1.16.

1.2.3 Validation Exercise

In this section I demonstrate that the drug confiscations data are a valid proxy for actual illicit drug flows. In particular, I correlate confiscations of imported drugs per capita (net of confiscations destined for other countries) in a locality to the availability of drugs in that locality. This approach is valid if local production is small relative to the local market, an assumption likely to hold in Spain as discussed in Section 1.2.1.

To measure local drug availability, I turn to the Survey on Alcohol and Drugs in Spain (EDADES). The EDADES is a nationally representative biennial survey on substance use in Spain, interviewing 20,000 to 30,000 persons per survey. Respondents are asked how easy it is for them to access various illegal drugs within 24 hours, how much of a problem illegal drugs are in their neighborhood, and whether they have personally used various drugs. I aggregate responses across the 2011, 2013, and 2015 survey rounds to create a measure of province-level drug use and drug availability.

I find that confiscations of illegal drugs positively correlate with a wide range of measures of local drug availability. In Figure 1.1, I plot the correlation coefficient between reported ease or difficulty obtaining a particular drug within 24 hours and the amount of that drug that was confiscated in the province per capita between 2011 and 2016.[19] Consistent with confiscations corresponding to real flows of illicit drugs, I find that when a higher proportion of respondents say it is "impossible" to obtain a particular drug, the amount of that drug confiscated in the province is lower. Conversely, I find that the proportion of respondents saying it is "easy" or "very easy" to obtain a drug correlates positively with the amount of that drug confiscated in the province. This relationship is much stronger for cannabis and cocaine, the major drugs imported into Spain, and weaker for heroin, whose pathway into Europe is generally believed to lie through the Balkan countries rather than through Spain (UNODC, 2014).

I also find that confiscations are weakly correlated with respondents' personal drug use history, as shown in Figure 1.17. I find a positive correlation between confiscations

[19] I do this exercise for cannabis, cocaine, and heroin, as respondents were not questioned about their access to amphetamines for the whole sample period. Respondents could reply that it was impossible, difficult, relatively easy, or easy to obtain the drug iwthin 24 hours.

and personal use for cocaine, with imprecise zeros for cannabis and heroin.

In Figure 1.2 I plot the correlation coefficients of various measures of local drug availability and use to the value of confiscations per capita across all illicit drugs. I measure local drug availability and use as the fraction of respondents replying that (in the first bar of Figure 1.2) drugs are a major problem in their neighborhood or that (for the remaining bars) they frequently see evidence of drug use and distribution in their neighborhood. For each survey question, confiscations vary positively with local drug availability.[20]

Overall, these results suggest that confiscations by law enforcement are a valid proxy for actual flows of illicit drugs. They are also consistent with Dobkin and Nicosia (2009), who find that drug markets quickly rebound even in response to confiscations of massive quantities of drugs.

1.3 Bilateral Empirical Analysis

I seek to understand whether immigrants facilitate drug trafficking between their origin country and their new home province. To do so, I relate drugs coming from a given origin country and confiscated locally with a measure of the number of immigrants from that origin country and living locally. Exploiting this country-province-pair level variation, I can flexibly control for observed and unobserved characteristics of the country and the province. Because migration and drug trafficking may be jointly determined by other factors, such as geographic or climatic similarity between country and province, I generate exogenous variation in the immigrant population using an instrumental variables strategy.

1.3.1 Preliminary Evidence

There exists a positive correlation between the number of immigrants and the value of drugs confiscated at the country-province level, as shown in Figure 1.18. This relationship may be driven by other factors, such as origin- or destination-specific institutions (e.g., economic development) or by country-province-pair factors such as geographic similarity. For example, consider the case of Morocco, a major source of both immigrants and cannabis flowing into Spain. Spatially, there is substantial overlap between the immigrant population and the location of confiscations of cannabis coming from Morocco (often on Spain's southern and eastern coast), as shown in Figure 1.3.

A natural explanation for this correlation is that geographic distance—since Morocco is directly to the south of Spain—drives both trafficking and immigration from Morocco and into southern Spain. Other confounders, such as the similar climate enjoyed by much

[20] Respondents are asked how often in their neighborhood they see people (i) drugged and on the ground, (ii) inhaling drugs in paper or aluminium, (iii) injecting drugs, (iv) selling drugs, (v) smoking joints, (vi) snorting drugs by nose, and (vii) leaving syringes lying on the ground.

of Spain and Morocco may also explain this correlation. To more formally evaluate the relationship between immigrants and drug trafficking and rule out such confounders, I next estimate a gravity equation of drug confiscations in the context of Spain.

1.3.2 Gravity Regression

My bilateral empirical specification, the gravity equation, allows me to control for origin- and destination-specific characteristics that may shape trafficking and migration. This estimation strategy also allows me to deal with concerns about enforcement intensity variation driving observed drug confiscations.

Specification. Given complete information on illegal drug flows, I would estimate a gravity equation of the form

$$\ln(X_{o,d}) = \alpha_o + \alpha_d + \beta M_{o,d} + \delta \ln(Dist_{o,d}) + \tilde{\varepsilon}_{o,d}$$

where α_o and α_d are origin and destination fixed effects, respectively, and $Dist_{o,d}$ is the distance in kilometers between o and d taken from Peri and Requena-Silvente (2010).[21] $M_{o,d}$ is a measure of the number of immigrants from o living in d, usually defined as the log of one plus the number of immigrants in d from o, measured in thousands (my results are robust to this functional form choice, as I show in Section 1.3.8). The error term $\tilde{\varepsilon}_{o,d}$ includes all omitted bilateral forces that may shape drug trafficking. I measure the immigrant population $M_{o,d}$ using the 2011 Spanish Census distributed by Minnesota Population Center (2019).

Because I cannot observe actual drug trafficking amounts, I instead use confiscations of illegal drugs. I denote the value of drugs confiscated in province d and coming from origin country o as $C_{o,d}$, where

$$C_{o,d} = E_{o,d} X_{o,d} \tag{1.1}$$

I define actual drug flows by value from origin country o to province d as $X_{o,d}$ and bilateral enforcement intensity as $E_{o,d} \in [0, 1]$, both of which are unobserved.

Plugging equation 1.1 into the gravity equation, I obtain my baseline specification,

$$\ln(C_{o,d}) = \alpha_o + \alpha_d + \beta M_{o,d} + \delta \ln(Dist_{o,d}) + \varepsilon_{o,d}$$

where $\varepsilon_{o,d} = \tilde{\varepsilon}_{o,d} + \ln(E_{o,d})$. The main parameter of interest is β, which measures the responsiveness of illegal drug confiscations to changes in the immigrant population.

The origin country and destination province fixed effects are key to my identification

[21] I provide microfoundations for this gravity equation in Appendix 1.6.

strategy. The origin fixed effect α_o controls for, among other factors, the economic development, institutions, and crime in the origin country as well as national-level policies of Spain vis-a-vis origin country o. These country-pair level policies can include visa regimes, customs regulations, and national law enforcement priorities. Similarly, the destination fixed effect α_d controls for province d factors common across origins, such as province d's police force strength and the economic conditions in d.

Thus β is identified off of variation in the drug confiscations and immigrant populations across country-province pairs. The identification assumption is that the country-province immigrant population $M_{o,d}$ is independent of country-province-specific enforcement intensity $E_{o,d}$ and any other country-province-level confounder $\tilde{\varepsilon}_{o,d}$.

For the empirical analysis, I replace the dependent variable $\ln C_{o,d}$ with $\ln(1 + C_{o,d})$ to avoid dropping bilateral links with no confiscations, as these make up more than half of my sample. I also estimate the immigrant-trafficking relationship using a dummy for whether any confiscation occurred as a dependent variable. Because drug confiscations are conducted locally, and therefore reporting practices may vary at the local level, I cluster standard errors at the province level.

In addition to imports, I explore how immigrants affect the re-exports of illegal drugs. Looking at both import and export margins allows me better understand the mechanisms underlying any immigrant-trafficking relationship. For example, if immigrants raise exports than immigrant demand for drugs is unlikely to drive the relationship. To measure intended re-exports, I consider drugs confiscated in d but that were intended to go to country o.[22] As dependent variables $Y_{o,d}$, I use either a dummy for whether any confiscation of drugs intended for re-export occurred $\mathbf{1}\{C_{d,o} > 0\}$ or the log of one plus the value of drugs confiscated and intended for re-export, $\ln(1 + C_{d,o})$.

For the main empirical analysis, my baseline gravity equation is

$$Y_{o,d} = \alpha_o + \alpha_d + \beta M_{o,d} + \delta \ln(Dist_{o,d}) + \varepsilon_{o,d} \tag{1.2}$$

OLS Results. In Table 1.1, I show OLS estimates when iteratively adding fixed effects controls. As expected, I find that including the province and country fixed effects significantly reduces the strength of the positive correlation between immigrants and drug confiscations. These estimates demonstrate the importance of including country and province fixed effects to reduce omitted variable bias, suggesting prior studies (Berlusconi et al., 2017; Giommoni et al., 2017; Aziani et al., 2019) may overstate the role of immigrants in facilitating drug trafficking.

[22] Note that I only observe confiscations of drugs entering Spain, so this measure excludes any drugs domestically produced for export.

1.3.3 Instrumental Variables Approach

While the country and province fixed effects absorb many potential confounders, there may still be unobserved factors at the country-province-pair level, such as the geographic or climatic similarity between a country and a Spanish province. To purge this potential endogeneity from country-province immigrant population, I adapt a leave-out push-pull instrumental variables approach to my setting. Consider, for example, that Moroccan immigrants settling in the province of Alicante may be drawn by its similar Mediterranean climate. Additionally, drug traffickers skilled at piloting boats in the waters off the coast of Morocco may be skilled at piloting boats in similar climates.

To obtain variation in migration exogenous to such concerns, I follow Burchardi et al. (2019) and develop a set of leave-out push-pull instruments for the number of immigrants arriving in a given region and coming from a given origin country. These instruments produce plausibly exogenous variation in bilateral immigrant inflows. I use two decades of inflows between 1991 and 2011 to predict the current number of immigrants from a given origin country living in a Spanish province.

The intuition of the instrument is that a social connection, in this case an immigration decision, between an origin and a destination is likely to occur when the origin is sending many immigrants at the same time the destination is pulling in many immigrants. For example, suppose we want to predict the number of Moroccans settling in the province of Alicante. To do so, we look at the number of Moroccans inflowing into Spain and the number of immigrants from all origin countries inflowing into Alicante for the same decade. In particular, the instrument will predict Moroccans to settle in Alicante if large numbers of immigrants from other countries are also settling there. Similarly, if many immigrants from other origins are settling in Alicante, then an immigrant arriving from Morocco will be predicted to settle in that province.

More specifically, the migration leave-out push-pull instrument interacts the arrival at the national level of immigrants from different origin countries (push) with the attractiveness of different destinations to immigrants (pull) measured by the fraction of immigrants settling in destination d. A simple version of the instrument predicts bilateral immigrant inflows and is defined as

$$\tilde{IV}^{D}_{o,d} = I^{D}_{o} \times \frac{I^{D}_{d}}{I^{D}}, \tag{1.3}$$

where I^D_o is the number of immigrants from origin o coming to Spain in decade D, I^D_d is the number of immigrants from all origins settling in destination province d in decade D, and I^D is the total number of immigrants arriving in Spain in decade D.[23]

However, the push-pull instrument defined in equation 1.3 may still fail the exclusion restriction. This may be the case if bilateral immigration is driven by endogenous con-

[23] An inflow from o to d is defined as a person interviewed in d for the 2001 or 2011 Spanish census with a nationality from o who arrived in the 10 years prior to the survey.

founders such as a similar climate in both origin and destination regions and if bilateral immigration is a large share of the instrument's individual components. Alternatively, there may be spatial correlation in confounding variables. For example, if both Moroccan and Algerian immigrants go to the province of Alicante due to the similar Mediterranean climates, then Moroccan migration to Alicante will be well predicted by Algerian migration so long as Algerian migration to Alicante is a sufficiently large share of total migration to Alicante. However, Algerian and Moroccan migration to Alicante may be jointly predicted by a third factor, climate, which may also affect drug trafficking (e.g., if calm weather facilitates smuggling by sea). To avoid such endogeneity, I again follow Burchardi et al. (2019) and leave out both the continent of origin country o and the autonomous community (the highest-level administrative unit in Spain) of province d to construct the instrumental variable defined as

$$IV_{o,d}^{D} = I_{o,-a(d)}^{D} \times \frac{I_{-c(o),d}^{D}}{I_{-c(o)}^{D}} \tag{1.4}$$

where $a(d)$ is the set of provinces in the autonomous community of d, and $c(o)$ is the set of countries on o's continent.

The identification assumption when using this instrument is that any confounding factors that make a given province more attractive for both immigration and drug trafficking from a given country do not simultaneously affect the interaction of (i) the settlement of immigrants from other continents with (ii) the total number of immigrants arriving from the same country but settling in a different autonomous community. A violation may occur if, suppose, immigrants skilled at drug trafficking from Morocco tend to settle in the province of Barcelona and immigrants skilled in drug trafficking from Lebanon settle in Alicante (Barcelona and Alicante are in different autonomous communities) in the same decade and for the same reason: a preference for the familiar Mediterranean climate. Moreover, if Moroccans are a large fraction of immigrants settling in Barcelona and Lebanese are a large fraction of the immigrants settling in Alicante, and therefore materially affect the instrument's prediction of flows of immigrants and drugs from Morocco to Alicante. Then the instrument is predicting bilateral immigration based on a confounding factor: climatic similarity between the immigrants' origin country and the Spanish province.

To measure immigrant inflows, I use the 2001 and 2011 Spanish Census from the National Institute of Statistics distributed by the Minnesota Population Center (2019). From these data, I use respondents' country of nationality, current province of residence in Spain, and year of migration. Since the set of origin countries for which I observe immigrant nationality differs for the two Census waves, I aggregate countries into the smallest consistent units allowable.

1.3.4 First-Stage

In Figure 1.4 I plot the first-stage fit of the instruments for the two decades of predicted inflows. The instruments vary positively with the log number of immigrants, as expected. Column 1 of Table 1.2 shows the first-stage regression coefficients. Instruments from both decades have a positive and statistically significant coefficient, and the first-stage F-statistic of 23.4 surpasses conventional threshold levels.

1.3.5 Results

I now turn to my baseline results on the effect of immigrants on illegal drug confiscations of imports and re-exports.

Table 1.2 shows the two-stage least squares estimation results for equation 1.2 for confiscations of imported drugs. Column 2 shows the results for the extensive margin of drug confiscations. The coefficient estimate of the effect of immigrants on the likelihood of a confiscation of imported illegal drugs for a country-province pair is 0.105 (SE = 0.039). This estimate implies that at the mean immigrant population at the province-country-pair level, 933, a 10% increase in the number of immigrants raises the likelihood that drugs trafficked from the origin country will be confiscated locally by 0.5 percentage points.[24] Similarly, in column 3, the coefficient estimate for the log of the immigrant population on the log value of illegal drugs confiscated is 2.33 (SE=0.56), which implies that a 10% increase in the immigrant population (again, at the province-country-pair level) relative to the mean raises the value of illegal drug imports confiscated by 12%.[25] This increase is in line with some estimates in the literature examining the effect of immigrants on *legal* trade.[26]

There are two biases relative to the OLS to take account of. First, there may be confounding variables at the country-province-pair level which drive both immigration and drug trafficking between locations. These confounders will tend to bias the OLS estimates upwards. Second, the number of immigrants from a given country living in a Spanish province may be mismeasured, biasing the OLS estimates downwards. My two-stage least squares estimates are statistically indistinguishable from the OLS estimates.

Table 1.3 shows the estimation results when the dependent variable is confiscations of drugs intended for re-export.[27] Column 2 shows the extensive margin result. The

[24]Using $\hat{\beta} = 0.105$ from column 2 in Table 1.2, can compute: $\mathbb{1}\left[C_{o,d}^{2011-2016} > 0 | M_{o,d}^{2011} = 933\right] = 0.105\left(\ln\left(1 + \frac{933\times1.1}{1000}\right) - \ln\left(1 + \frac{933}{1000}\right)\right) \approx 0.0049$.

[25]Using $\hat{\beta} = 2.331$ from column 3 in Table 1.2, we have: $\frac{C_{o,d}^{2011-2016}[M_{o,d}^{2011}=1.1\times933]}{C_{o,d}^{2011-2016}[M_{o,d}^{2011}=933]} - 1 = \exp\left(2.331\left(\ln\left(1 + \frac{1.1\times933}{1000}\right) - \ln\left(1 + \frac{933}{1000}\right)\right)\right) - 1 = 0.116$.

[26]See, for example, Parsons and Vézina (2018), who estimate the effect of a 10% increase in immigrant population raises the value of legal trade by 4.5% to 13.8%.

[27]Column 1 restates the first-stage estimates.

coefficient estimate for the effect of immigrants on the likelihood of drugs imported from a given origin country and confiscated locally is 0.083 (SE = 0.021). This estimate implies that at the mean immigrant population, 933, a 10% increase in the number of immigrants raises the likelihood that the link will be used for drug trafficking by 0.4 percentage points.[28] Similarly, in column 3, the coefficient estimate of the log of the immigrant population on the log value of drugs confiscated is 1.339 (SE=0.34), which implies that a 10% increase in the immigrant population relative to the mean raises the value of drug imports confiscated by 6.5%.[29]

1.3.6 Preferences for Drugs and Trade Costs

After controlling for the institutions and labor market conditions of the host province and origin country, more immigrants may raise imports of illicit drugs for two reasons. First, they may prefer to consume goods imported from their home country. Second, immigrants reduce trade costs between origin and destination.

Immigrant Preferences. Atkin (2013) and Bronnenberg et al. (2012) suggest that immigrants may share the same tastes for food and other products as consumers in their origin country. If these similar tastes also apply to illicit drugs, more drugs may be trafficked from immigrants' origin country. However, such a story would require retail drug consumers to have an implausible combination of tastes and information. Consider an immigrant from Venezuela who consumes cocaine. This immigrant would need to be able to distinguish street cocaine based on which country it was trafficked from (not produced in). However, since the modifications to cocaine generally occur close to the point of production and in any case do not differ much based on production location, it is unlikely that the immigrant's experience would differ much based on which country the cocaine was trafficked through.

I also compare drug use between immigrants and native-born Spaniards and find that immigrants consume drugs at a substantially lower rate. Using the EDADES data introduced in Section 1.2.3 for the years 2005 through 2015, I find that 22% of those born outside of Spain have ever consumed cannabis, cocaine, heroin, or amphetamines compared to nearly 35% of native-born Spaniards. Taken together, these facts suggest immigrants bringing the demand for drugs from their home country with them to Spain are unlikely to explain my baseline results.

[28] Using $\hat{\beta} = 0.083$ from column 2 in Table 1.3, can compute: $\mathbb{1}\left[C_{d,o}^{2011-2016} > 0 | M_{o,d}^{2011} = 933\right] = 0.083\left(\ln\left(1 + \frac{933\times 1.1}{1000}\right) - \ln\left(1 + \frac{933}{1000}\right)\right) \approx 0.0039$.

[29] Using $\hat{\beta} = 1.339$ from column 3 in Table 1.3, we have: $\frac{C_{d,o}^{2011-2016}[M_{o,d}^{2011}=1.1\times 933]}{C_{d,o}^{2011-2016}[M_{o,d}^{2011}=933]} - 1 = \exp\left(1.339\left(\ln\left(1 + \frac{1.1\times 933}{1000}\right) - \ln\left(1 + \frac{933}{1000}\right)\right)\right) - 1 = 0.065$.

Trade Costs. Immigrants may increase illegal trade in much the same way they raise legal trade. Felbermayr et al. (2015) note that immigrant networks can reduce information and search frictions for trade between two locations, since trust may be greater within nationality and information travels more smoothly within nationality group. Additionally, immigrant networks raise the cost of opportunistic or cheating behavior by firms within the nationality network, who can be punished for bad behavior by being shunned from business within the network (Rauch and Trindade, 2002). Finally, the qualitative studies summarized in Section 1.2.1 demonstrate ways in which social connections between immigrants can facilitate trafficking by reducing trade costs.

In the context of this study, I find that immigrants raise drug flows on both the import and re-export margin. The fact that immigrants increase re-exports suggest that immigrants reduce trade costs rather than simply raise demand for drugs.

1.3.7 Drug-Hub Level of Immigrant's Origin Country

To understand the degree to which the immigrant-trafficking relationship is heterogenous by origin country, I look at whether drugs being confiscated are coming from countries that are hubs of drug trafficking.[30] I re-estimate equation 1.2, interacting the country-province immigrant log population with the drug-hub level of the immigrants' origin country.[31]

In Table 1.11 I show the estimated coefficients. I find that origin countries that are significantly involved in drug trafficking, that is, send a substantial amount of illicit drugs to countries other than Spain, are more likely to export drugs to Spain when more immigrants from those countries settle in Spain.

1.3.8 Robustness Checks and Legal Trade

In my baseline analysis, I make specific assumptions on my functional form, sample, and specification. Below, I show that my baseline results are robust to variations on each of these dimensions.

Relaxing Functional Form Assumption

In my baseline specification, equation 1.2, I measure the endogenous variable of interest as the log of one plus the number of immigrants measured in thousands, $\ln\left(1+\frac{migrants_{o,d}^{2011}}{1000}\right)$.

[30] I define the drug-hub level of a given country as either the fraction of global drug confiscations for which the country was the exporter or the rank order thereof.

[31] Data on world bilateral drug confiscations are similarly taken from the UNODC dataset on individual drug confiscations that I use for Spain. One drawback of these data for countries other than Spain is that reporting of drug confiscations to the UNODC occurs less frequently and is of lower quality. Nevertheless, no alternative data source on country-pair drug trafficking exists, so I pursue this analysis using these imperfect data.

To test whether my results are sensitive to changes in the function form of the endogenous variable, I perform several robustness exercises.

First, I estimate my baseline specification across a range of alternative functional forms for the number of immigrants, including a linear term for the immigrant population. I show the results in Tables 1.9 and columns 1–3 of 1.10. Across functional forms, more immigrants still lead to more drug confiscations. In addition, I estimate 1.2 using Poisson Pseudo Maximum Likelihood (PPML) (Silva and Tenreyro, 2006, 2011). PPML estimation has the advantage of allowing for zeros in the dependent variable without introducing potential distortion by adding one within the log function. As shown in column 4 of 1.10, I find that PPML estimation still generates a positive relationship between immigrants and the value of illegal drugs confiscated.

Next, I relax the log-functional form assumption.[32] Specifically, I estimate

$$\begin{aligned}\mathbf{1}\left[C_{o,d}>0\right] &= \delta_o+\delta_d+\beta_1\ln(1+\pi_1 migrants^{2011}_{o,d})+\epsilon_{o,d}\\ \ln(C_{o,d}) &= \alpha_o+\alpha_d+\beta_2\ln(1+\pi_2 migrants^{2011}_{o,d})+\varepsilon_{o,d}\end{aligned} \tag{1.5}$$

In equation 1.5 I estimate (π_1,π_2) whereas in equation 1.2 I assume $\pi_1=\pi_2=0.001$. I do so using non-linear generalized method of moments using moment conditions

$$E\left[\boldsymbol{Z_{o,d}}\times(Y_{o,d}-\alpha_o-\alpha_d-\beta_1\ln(\pi_1 migrants^{2011}_{o,d}+1))\right]=\mathbf{0}$$

$$E\left[\begin{pmatrix}\boldsymbol{\alpha_o}\\ \boldsymbol{\alpha_d}\end{pmatrix}\times(Y_{o,d}-\alpha_o-\alpha_d-\beta_1\ln(\pi_2 migrants^{2011}_{o,d}+1))\right]=\mathbf{0}$$

for dependent variable $Y_{o,d}\in\{\ln(C_{o,d}+1),\mathbf{1}\left[C_{o,d}>0\right]\}$ and instrument set

$$\boldsymbol{Z_{o,d}}=\left(I^{IV,1991-2001}_{o,d},I^{IV,2001-2011}_{o,d},(I^{IV,1991-2001}_{o,d})^2,(I^{IV,2001-2011}_{o,d})^2\right)'$$

I include squared terms for the instruments to improve convergence and add a moment for the constant, thus yielding 163 moments. Similar to my baseline estimation, I cluster standard errors by province.

Table 1.4 shows the results. My estimates of (π_1,π_2), do not reject my baseline functional form assumption of $\pi_1=\pi_2=\frac{1}{1000}$ and reject the more conventional functional form choice $\pi_1=\pi_2=1$. In addition, the estimates of (β_1,β_2) also are statistically indistinguishable from my baseline coefficient estimates. At the point estimates, I find that a 10% increase in the number of immigrants relative to the mean raises the probability of

[32] I motivate my choice of a log-functional form with the binscatter plot in Figure 1.24 of the relationship between the immigrant population and the dummy variable for whether any confiscation occurs at the country-province level.

a confiscation occurring on a bilateral link by 1.1 percentage points and raises the value of illegal drugs confiscated by 20%.

Finally, I relax completely my functional form assumption by estimating a non-parametric regression relating import drug confiscations to the number of immigrants following Chetverikov and Wilhelm (2017). Figure 1.21 depicts the results. While I find a weakly increasing relationship between immigrants and import drug confiscations, the standard errors are very large. Nevertheless, I take this as suggestive evidence supporting the baseline parametric estimation results.

Varying Estimation Sample

Drug trafficking into Spain is primarily driven by a select few countries—Morocco, for example, is the dominant exporter of cannabis to Spain. To see whether any particular origin country drives my baseline results, I re-estimate the gravity specification, leaving out individual countries. Figure 1.22 shows the distribution of β estimates from equation 1.2 when I drop one origin country at a time for both dependent variables, $\mathbf{1}\left[C_{o,d} > 0\right]$ and $\ln(C_{o,d} + 1)$. The histograms show that I estimate a positive β regardless of which country I drop from the sample, suggesting that no single country drives the results.

I also estimate the immigrant-confiscations relationship separately by drug type. For cannabis and cocaine, I estimate positive and statistically significant effect sizes. Cocaine appears to be more reliant on immigrants for importation than cannabis, which can be produced locally in contrast to cocaine, which must be imported. For heroin and amphetamines, the effect is close to zero, as shown in Figure 1.23. However, heroin and amphetamines represent less than 1% of drugs confiscated by Spain, as shown in Figure 1.11 and therefore precise estimates are difficult to obtain.

Finally, I consider a selection of high-trafficking countries and provinces alone. In Figure 1.19, I show the relationship between import drug confiscations and immigrants graphically for Morocco and Colombia and two of the largest receiving provinces, Madrid and Barcelona. In Figure 1.20, I do this for re-exports with France and Italy and again with Madrid and Barcelona. In every case, more immigrants lead to more confiscations.

Standard Errors

In my baseline specification, I cluster standard errors at the province level, as this is the level of police reporting of confiscation events. Table 1.12 shows estimates using different clustering of standard errors, and they mostly remain statistically significant across the different clustering geographies.

Panel Estimation

In my baseline cross-sectional estimation, I argued that the push-pull instrumental variable dealt with country-province-pair level confounders, such as the similar climate of the country and province. To gauge the extent to which the instrument takes care of such time-invariant country-province endogeneity, I estimate a panel specification, specifically

$$Y_{o,d,t} = \alpha_{o,t} + \alpha_{d,t} + \delta \ln(Dist_{o,d}) + \beta M_{o,d,t} + \varepsilon_{o,d,t} \tag{1.6}$$

$$Y_{o,d,t} = \alpha_{o,t} + \alpha_{d,t} + \alpha_{o,d} + \beta M_{o,d,t} + \varepsilon_{o,d,t} \tag{1.7}$$

where $Y_{o,d,t} \in \{\mathbf{1}[C_{o,d,t} > 0], \ln(C_{o,d,t} + 1)\}$ for the value of drugs confiscated in d from o in year t $C_{o,d,t}$ for both imports and intended re-exports. $M_{o,d,t}$ is measured as the log of the bilateral immigrant population in thousands plus 1, where the bilateral immigrant population is derived from annual tabulations taken from Spain's local population registries at the country-by-province level. I estimate equations 1.6 and 1.7 for the years 2002 through 2016.

I modify the instrumental variables for the panel analysis by including the cross-sectional 1991–2001 push-pull instrument

$$IV_{o,d}^{1991-2001} = I_{o,-a(d)}^{1991-2001} \times \frac{I_{-c(o),d}^{1991-2001}}{I_{-c(o)}^{1991-2001}} \tag{1.8}$$

as well as a time-varying instrument that predicts bilateral immigrant inflows between 2001 and year t,

$$IV_{o,d,t}^{recent\ \ years} = I_{o,-a(d)}^{2001-t} \times \frac{I_{-c(o),d}^{2001-t}}{I_{-c(o)}^{2001-t}} \tag{1.9}$$

I compute immigrant inflows between 2001 and t as the net change in the bilateral immigrant population as measured in the population registry. To improve the first-stage fit (and similar to Burchardi et al., 2019), I also add squared versions of the instrumental variables.[33]

For imports, I estimate equation 1.6 in Table 1.13 and equation 1.7 in Table 1.14. For re-exports, I estimate equation 1.6 in Table 1.15 and equation 1.7 in Table 1.16. The estimated coefficients are in line with my baseline estimates in Tables 1.2 and 1.3, suggesting that time-invariant country-province level confounders are not significantly shaping my baseline results.

[33] Without the squared terms, I obtain a first-stage F-statistic of approximately 14. My second stage results also carry through without the squared terms for the instrumental variables.

Legal Trade

To gauge the magnitude of the effect size estimated in Section 1.3.5 for illegal trade relative to legal trade, I estimate the relationship between the bilateral immigrant population and legal trade. To measure legal trade volume, I turn to the ADUANAS-AEAT dataset provided by the Spanish government. This dataset provides transaction-level data and includes information on the origin (for imports) or destination (for exports) country and the same for the origin or destination province within Spain. I aggregate these data to the province-by-origin country level for imports for the years 2011 to 2016.

Because I find some sensitivity of this relationship with respect to functional form choices, I estimate the generalized method of moments with moments

$$E\left[(\ln(X_{o,d}^{legal}+1)-\delta_2-\beta_2\ln(1+\pi_2 migrants_{o,d}^{2011}))\times Z_{o,d}\right]=0$$

where $X_{o,d}$ is the value of legal goods imported into province d originating from country o and for instrument set

$Z_{o,d}=\left(I_{o,d}^{IV,91-01}, I_{o,d}^{IV,01-11}, (I_{o,d}^{IV,91-01})^2, (I_{o,d}^{IV,01-11})^2, (I_{o,d}^{IV,91-01}\times I_{o,d}^{IV,01-11})\right)'$.[34]

Column 2 of Table 1.4 shows the results. I estimate that a 10% rise in the number of immigrants increases legal trade by about 13%, a magnitude comparable to the effect of immigrants on illegal drug confiscations.[35]

1.4 General Equilibrium Responses and Enforcement Intensity

My gravity estimates may not imply that overall illegal drug market activity rises with additional migration for two reasons. First, increases in the bilateral immigrant population may increase the scrutiny of law enforcement, thus resulting in the relationship estimated in Section 1.3.5 but not corresponding to a real rise in actual drug flows. Second, increases in trafficking may be offset by decreases in local production or decreases in imports on other bilateral links. I do not find evidence for either of these channels, as I show below.

[34] With nearly every province-origin country pair having positive trade I do not have enough variation along the extensive margin of trade to also estimate the comparable moment for legal trade.

[35] As shown in column 2 of Table 1.4, I estimate that $\hat{\beta}=1.36$, $SE=0.1$ and $\hat{\pi}=0.013$, $SE=0.0068$. To get the elasticity from this nonlinear equation, I compute that $\frac{X_{o,d}^{2011}[M_{o,d}^{2011}=1.1\times 963]}{X_{o,d}^{2011}[M_{o,d}^{2011}=963]}-1=\exp\left(1.36\left(\ln\left(1+0.012\times(1.1\times 963)\right)-\ln\left(1+0.012\times 963\right)\right)\right)-1=0.127$.

1.4.1 Enforcement Intensity

In Section 1.2.3 I showed that drug confiscations correspond to drug use and availability at the *province level*. In my bilateral estimation, I control for enforcement intensity specific to each Spanish province (and common across all origins) as well as for enforcement intensity specific to each origin country (but common to all Spanish provinces). In this section, I conduct two exercises at the *bilateral level* to assess the extent to which variation in bilateral enforcement intensity drives my baseline results from Section 1.3. I also conduct an additional test for the extent to which enforcement intensity drives confiscations in Appendix 1.6.

Quantitative Exercise

First, I consider the plausibility of variation of enforcement intensity explaining the quantitative magnitudes that I estimated in Section 1.3.2. In particular, I ask how much bilateral enforcement intensity would have to increase to fully explain the observed effect of immigrants on drug confiscations.

To formalize this notion, take the derivative of equation 1.1 with respect to the number of immigrants:

$$\frac{dC_{o,d}}{dM_{o,d}} = E_{o,d}\frac{\partial X_{o,d}}{\partial M_{o,d}} + X_{o,d}\frac{\partial E_{o,d}}{\partial M_{o,d}} \tag{1.10}$$

Dividing equation 1.10 by the value of drugs confiscated $C_{o,d}$ and multiplying by the immigrant population $M_{o,d}$, I obtain

$$\epsilon_{C,M} = \epsilon_{X,M} + \epsilon_{E,M} \tag{1.11}$$

where $\epsilon_{a,b}$ is the elasticity of a with respect to b. In Section 1.3.2, I estimate $\hat{\epsilon}_{C,M} = 1.2$. Suppose now that actual drug flows are not at all affected by the bilateral immigrant population, that is, $\epsilon_{X,M} = 0$. To assess the plausibility of this assumption, I first calculate a back-of-the-envelope estimate of the elasticity of enforcement intensity to immigrant population, $\hat{\epsilon}_{E,M}$.

I consider the effects of a 2 standard deviation increase in the predicted bilateral immigrant population, residualized on origin and destination fixed effects and log distance. The median of predicted immigrants is 11, and a 2 standard deviation increase raises this to 332.[36] This represents an increase in the country-province-specific immigrant population of 3000%, which would require a 3600% increase in enforcement intensity if my results were driven entirely by changes in enforcement.

[36] Where $11 \approx (exp(0.11) - 1) \times 1000$ and the standard deviation of residualized bilateral immigrant population is ≈ 0.14.

To gauge the size of the implied increase in enforcement intensity, I compute a rough estimate of the fraction of drugs confiscated by Spain. I calculate this as

$$\hat{E}_{Spain} = \frac{C_{Spain}}{Y_{EU} \times \frac{C_{Spain}}{C_{EU}} + C_{Spain}}$$

where Y_{EU} is the size of the market for illegal drugs in the European Union and C_X is the value of drugs confiscated by X. I focus on the market for cannabis and cocaine, as they are the primary drugs appearing in the Spanish confiscations data.

For Y_{EU}, I use the European Monitoring Centre for Drugs and Drug Addiction[37] estimate for the size of the market for cocaine and cannabis in the European Union of about 20 billion USD in 2013. I compute $\frac{C_{Spain}}{C_{EU}}$ using the international UNODC confiscations data and find that Spain confiscated 78% of cannabis and cocaine by value. Between 2011 and 2016, on average 1 billion USD worth of cocaine and cannabis was confiscated by Spanish authorities each year. I therefore compute that about 6% of cocaine and cannabis entering Spain are confiscated by Spanish law enforcement. Therefore an increase in enforcement intensity of 3600% would raise enforcement intensity to 2.17, which is infeasible since $E_{o,d} \leq 1$.

Extensive Margin of Trafficking

Next, I use the intuition that for bilateral links near the extensive margin of trafficking drugs, enforcement changes caused by variation in the number of immigrants will not be important in driving confiscations.

In my baseline estimation, I assume that $\frac{\partial E_{o,d}}{\partial M_{o,d}} = 0$ in equation 1.10, allowing me to estimate the object of interest, $\frac{\partial X_{o,d}}{\partial M_{o,d}}$. However, my estimation will also pick up changes in bilateral enforcement intensity that result from changes in bilateral migration, $\frac{\partial E_{o,d}}{\partial M_{o,d}}$. This may occur if, for example, police target immigrant groups for drug trafficking enforcement actions once that group reaches a critical mass.

To test this assumption and gauge the extent to which enforcement intensity variation may affect my results, I estimate

$$\mathbf{1}\{C_{o,d} > 0\} = \alpha_o + \alpha_d + \beta M_{o,d} + \delta \ln(Dist_{o,d}) + \varepsilon_{o,d} \tag{1.12}$$

for the subset of observations for which I predict that $X_{o,d} \approx 0$.[38]

To predict when actual flows $X_{od} \approx 0$, I use a similar leave-out push-pull structure for confiscations as I did for immigrant inflows:

[37] https://www.emcdda.europa.eu/system/files/publications/3096/Estimating%20the%20size%20of%20main%20drug%20markets.pdf

[38] Akee et al. (2014) similarly focus on the extensive margin when estimating the determinants of transnational human trafficking.

$$\hat{C}_{o,d} = C_{o,-a(d)} \times \frac{C_{-c(o),d}}{C_{-c(o)}} \tag{1.13}$$

where $\hat{C}_{o,d}$ interacts confiscations of drugs originating from o but confiscated outside the autonomous community of d with the fraction of all drugs from outside o's continent confiscated in d. Implicit in this formulation is the assumption that (1) on average, law enforcement in province d will discriminate differently against immigrants from continents outside of $c(o)$, and (2) on average, law enforcement in other autonomous communities will discriminate differently against immigrants from o.

I show results in Table 1.5 subsetting to bilateral links that I predict having less than \$1,000 worth of drugs confiscated. While the point estimate falls when subsetting to the sample predicted to be on the extensive margin, the two estimates in columns 1 and 2 are statistically indistinguishable, suggesting enforcement variation cannot fully explain my bilateral results.

1.4.2 General Equilibrium Responses

While I have shown that more immigrants on a bilateral link raise bilateral drug confiscations, this effect may be offset by general equilibrium adjustments to immigrant-induced trafficking. For example, immigrants from one country may adjust their trafficking in response to more immigration from another country. If such adjustments offset the effect of immigrants on trafficking, then there should be no effect when aggregating across origin countries. To assess the strength of the general equilibrium response, I estimate the effect of immigrants on drug market activity at the province level.

Drug Confiscations and Use

I first estimate the effect of immigrants on confiscations of illegal drugs and illegal drug use with a panel of Spanish provinces. For the years 2003 to 2016, I estimate

$$\ln Y_{d,t} = \alpha_d + \alpha_t + \beta \ln M_{d,t} + \epsilon_{d,t} \tag{1.14}$$

for some measure Y_d of illegal drug activity in d and the log number of immigrants from all origins $M_{d,t}$ in year t. I also control for province and year fixed effects and cluster standard errors at the autonomous community-by-year level. Because there might be factors affecting both immigration and drug smuggling into a province, I instrument for the immigrant population using the shift-share instrumental variable from Cortes (2008):

$$IV_{d,t} = \ln \left[\sum_o \left(\frac{Immigrants_{o,d,1981}}{Immigrants_{o,1981}} \right) \times Immigrants_{o,t} \right] \tag{1.15}$$

where $Immigrants_{o,t}$ refers to the number of immigrants from o living in Spain in year t.[39]

Because I am exploiting less variation than in my baseline gravity estimation, interpreting β as the causal effect of immigrant share on drug activity requires a stronger identifying assumption, as I can no longer exploit variation across immigrant origins. In particular, my identification assumption requires that there are no persistent shocks within autonomous communities that shape the distribution of immigrant populations in 1981, the distribution of immigrant populations in in the 2000s, and the distribution of drug trafficking across space in the 2000s.

In Figure 1.25 I show the first-stage fit. The instrument well predicts the immigrant population across Spanish provinces over time.

I estimate equation 1.14 with dependent variable $C_{d,t}$, the log value of drugs confiscated in province d in year t. Column 2 of Table 1.6 shows the result. I find that a 1% increase in immigrant population share in a province raises drug smuggling into that province by 19% overall. This elasticity of immigrant population to illegal drugs imported is higher than my baseline estimates, suggesting general equilibrium adjustment (such as trade diversion) to trafficking by immigrants does not offset the effect of immigrants on trafficking.

I next estimate equation 1.14 with dependent variable $DrugUsers_{d,t}^{Native}$, the number of native-born drug users per capita measured using the EDADES survey described in Section 1.2.3. I find no effect of immigrants on the drug use of the native-born as shown in columns 3 and 4 of Table 1.6, perhaps because immigrant-induced drug trafficking is mostly re-exported, and is therefore not intended for use in the local market.

Drug Arrests and Cultivation

Next, I estimate the effect of the immigrant population on arrests for drug trafficking and domestic cultivation of cannabis. Due to a lack of data, I use a single cross-section of Spanish provinces. I estimate

$$\ln Y_d = \alpha + \beta \ln M_{d,2011} + \gamma \ln Population_{d,1981} + \epsilon_{d,t} \tag{1.16}$$

for some measure Y_d of illegal drug activity in d and the log number of immigrants from all origins $M_{d,2011}$ in 2011. I again use the shift-share instrumental variable defined in equation 1.15.

I first estimate 1.16, measuring illegal drug activity Y_d as the number of native-born Spaniards arrested for drug trafficking offenses in province d between 2011 and 2016. I

[39] I also use a jackknife version of equation 1.15 in which I leave out province d, that is $IV_{d,t} = \ln\left[\sum_o \left(\frac{Immigrants_{o,d,1981}}{Immigrants_{o,1981}}\right) \times Immigrants_{o,-d,t}\right]$. I show results in Table 1.17.

find that a larger immigrant population does not lead to statistically significant differences in drug trafficking arrest rates of the native-born, as shown in column 2 of Table 1.7.

Finally, I measure illegal drug activity as the log number of cannabis plants confiscated. Spain produces a small but non-trivial amount of cannabis.[40] I draw on Alvarez et al. (2016), who assemble a dataset on cannabis plant confiscations based on 2013 press reports and public statements by the Spanish government.[41] I find that as the local immigrant population increases, there is no effect on the number of cannabis plants confiscated locally, suggesting there is not a large domestic production response to changes in immigrant drug trafficking.

1.5 Legal Status, Naturalization, and Trafficking

Immigrants' integration into labor markets and civil society may be hampered when they do not have legal status or a path to citizenship. A lack of legal status may hinder their access to the formal labor market, which lowers the opportunity cost of crime (Becker, 1968; Ehrlich, 1973). This may result in an increase in criminal activity among immigrants, as found empirically by Mastrobuoni and Pinotti (2015), Pinotti (2017), and Freedman et al. (2018). Hence a lack of legal status may lead immigrants to illegally traffic drugs.

To assess whether this intuition holds for drug trafficking, I conduct two exercises. First, I use a gravity equation to estimate separately the effect of irregular immigrants (those without legal status) and regular immigrants on drug confiscations and find that my bilateral results are driven entirely by irregular immigrants. Second, I exploit an extraordinary regularization program in 2005 to explore the long-run dynamics of receiving legal status and later obtaining citizenship, and I find that granting immigrants citizenship can significantly reduce drug trafficking.

1.5.1 Measuring the Irregular Immigrant Population

To estimate the prevalence of irregular immigrants at the origin country-destination province level, I take the difference between the number of persons appearing in the population registry of province d from origin country o and the number of persons with

[40] Alvarez et al. (2016) find that in 2013, authorities confiscated almost 200,000 cannabis plants growing in Spain. Combining the United Nations' estimate of the average weight of a cannabis plant (p. 39, UNODC, 2017) with the estimate of wholesale prices of cannabis herb in Spain for 2013, the confiscated plants are valued at approximately \$26 million. This compares to about \$312 million in confiscated cannabis coming from outside Spain.

[41] I do not have access to the microdata compiled by Alvarez et al. (2016), but instead use the approximate number of plants confiscated by province derived from their Figure 4. This leads to some measurement error. Moreover, I do not observe confiscations in the provinces of Ceuta or Mellila.

residency permits in province d from country o. Specifically, I compute

$$Irregular\ \ Migrants_{od} = Population\ \ Registry\ \ Count_{od} - Residency\ \ Permits_{od} \tag{1.17}$$

and then divide by the total bilateral immigrant population to obtain the fraction of immigrants who have irregular status.

I do this for all 52 Spanish provinces as well as for the 75 origin countries for which I observe bilateral population registry figures and bilateral residency permits in 2011. I estimate that 27% of immigrants living in Spain are irregular, consistent with the estimate from González-Enríquez (2009) in 2008.

1.5.2 Gravity Estimation by Legal Status

To explore whether irregular migration is an important factor in explaining the connection I find between immigrants and drug trafficking, I modify my baseline specification, equation 1.2, to include two separate terms for the bilateral immigrant population by regular ($M_{o,d}^{reg}$) and irregular ($M_{o,d}^{irreg}$) status:

$$Y_{o,d} = \alpha_o + \alpha_d + \beta_{irreg} M_{o,d}^{irreg} + \beta_{reg} M_{o,d}^{reg} + \zeta \ln(Dist_{o,d}) + \varepsilon_{o,d} \tag{1.18}$$

where, as in the baseline, $Y_{o,d}$ is either a dummy for whether any confiscation occurred or the log value of drugs confiscated plus one. Thus β_{irreg} is the effect of irregular immigrants on trafficking and β_{reg} is the effect of regular immigrants on trafficking.

Separating immigrants by legal status introduces another endogeneity issue—differential selection of immigrants into legal status and trafficking—which the baseline leave-out push-pull instrument defined in equation 1.4 may not address. In particular, there may be some characteristic of immigrants, such as a taste for risk-taking, which drives selection into both irregularity and drug trafficking. To the extent that this selection is common across provinces for a given nationality, the country fixed effect α_o will absorb such selection. Similarly, if the characteristic is common across immigrants of different nationalities in a given province, the province fixed effect α_d will absorb this.

To address province-country-specific selection into irregularity and drug trafficking, I modify the leave-out push-pull instrument predicting immigrant inflows to predict immigrant inflows by legal status. In particular, I interact the leave-out push-pull instrument with the lagged leave-out fraction of immigrants with legal status L,

$$IV_{o,d}^{D,L} = m_{o,d}^{L} \times IV_{o,d}^{D} \tag{1.19}$$

for $L \in \{regular, irregular\}$ and decade D, where $m_{o,d}^{L} = \frac{immigrants_{o,-a(d)}^{2003,L}}{immigrants_{o,-a(d)}^{2003}}$, the fraction of immigrants with legal status L from country o who live outside the autonomous community of province d back in 2003. The instrument interacts variation across three dimensions: (i) immigration from various origin countries, (ii) immigration to various Spanish provinces, and (iii) the propensity of immigrants to have legal status L at the country-province level. The identification restriction is that there are no confounders—either persistent from 2003 to 2011 or present in both province d and another province outside d's autonomous community—at the province-country-pair level driving selection of immigrants into both irregular status and drug trafficking.

More concretely, consider the case of Moroccan immigrants living in Barcelona. $m^L_{o,d}$ uses information on the legal status of Moroccan immigrants outside Catalonia (the autonomous community of Barcelona) back in 2003 to predict the 2011 legal status of Moroccans in Barcelona. The exclusion restriction is violated if, for example, Moroccans in Madrid in 2003 were driven into irregularity and drug trafficking by the same confounder (e.g., a preference for risk-taking) that drove Moroccans in Barcelona in 2011 into irregularity and trafficking—so long as a non-trivial share of Moroccans outside Catalonia live in Madrid and the confounder acts disproportionately on Moroccans in Madrid than on Moroccans elsewhere (i.e., it is not absorbed by the Moroccan fixed effect).

In Panel A of Table 1.8 I show the results for estimating equation 1.18; in Panel B I show results when using the instruments defined in equation 1.19. I find that a 10% increase in the bilateral *irregular* immigrant population raises the likelihood of an illegal drug confiscation by 1.9 percentage points (column 2 of panel B). By contrast, a 10% increase in the bilateral *regular* immigrant population slightly reduces illegal drug confiscations and the estimated coefficient is statistically insignificant.[42] A 10% increase in the bilateral population of irregular immigrants raises the value of drugs confiscated by 29%, while 10% increase in the bilateral regular immigrant population leads to a small and statistically insignificant increase in the value of drugs confiscated (column 4 of panel B).[43] Effect sizes may be larger in the 2SLS (panel B) than in the OLS (panel A) due to greater measurement error in the population of immigrants by legal status, since I estimate these populations as outlined in Section 1.5.1.

These results suggest immigrant legal status is an important factor shaping immigrants' role in drug trafficking. However, the composition of immigrants for a given country-province pair may differ based on the immigrants' legal status. To better understand the role immigration policy can play in mitigating the immigrant-trafficking relationship, I turn to an event study of a major immigrant regularization.

---

[42]Using $\hat{\beta}^{Reg} = -0.112$ from column 2 and mean value of bilateral immigrant population of 933, I find that $\mathbb{1}\left[C^{2011-2016}_{o,d} > 0 | M^{2011}_{o,d} = 933\right] = -0.112\left(\ln\left(1 + \frac{933\times 1.1}{1000}\right) - \ln\left(1 + \frac{933}{1000}\right)\right) \approx 0.005$ and for $\hat{\beta}^{Irreg} = 0.403$, this is 0.019.

[43]Using $\hat{\beta}^{Reg} = 0.0383$ from column 4, we have: $\frac{C^{2011-2016}_{od}[M^{2011}_{o,d}=1.1\times 933]}{C^{2011-2016}_{od}[M^{2011}_{o,d}=933]} - 1 = \exp\left(0.0383\left(\ln\left(1 + \frac{1.1\times 933}{1000}\right) - \ln\left(1 + \frac{933}{1000}\right)\right)\right) - 1 \approx 0.0018$. For irregular migration, this is 0.29.

1.5.3 2005 Mass Regularization Event Study

In 2005, Spain conducted the largest regularization event of immigrants in its history, with over half a million immigrants obtaining legal status. Immigrants who were registered with their local council in the population registry as of August 8, 2004, were offered a work contract of at least six months (three months if in agriculture), and had no criminal record in their home country or in Spain, were eligible to apply for regular status, usually through their prospective employer (González-Enríquez, 2009).

To better understand the effects of the regularization, I estimate an event study at the province-by-year level. This differs from my baseline cross-section estimates in Section 1.3.2 in that I use year-to-year variation in drug confiscations. At the bilateral level, confiscations can occur highly irregularly, with no confiscations for several years followed by a year with one massive confiscation. This is likely more a result of variation in enforcement "luck" rather than changes in actual flows of illicit drugs, and therefore it reflects measurement error. To smooth out this variation and thereby obtain more precise estimates, I aggregate to the province level. Doing this has the added benefit of improving measurement of the number of irregular immigrants, as the bilateral-level measurement excludes many countries and appears to censor bilateral links with very few immigrants.

I estimate this event study using the equation

$$Y_{d,t} = \sum_{t \neq 2004} \theta_t \times m_d^{2003,irregular} + \delta_d + \delta_t + \epsilon_{d,t} \tag{1.20}$$

where $m_d^{2003,irregular}$ is the number of irregular immigrants in 2003 imputed as in equation 1.17. I plot the θ_t coefficients in Figure 1.5, both for whether any confiscation occurred, $Y_{d,t} = \mathbf{1}\{C_{d,t} > 0\}$, and the log value of drug confiscations, $Y_{d,t} = \ln(C_{d,t} + 1)$.

I find that the 2005 regularization led to a sudden jump in the number of work authorizations granted to immigrants in Spain, as shown in Figure 1.26. In addition, naturalizations of immigrants increased markedly in 2005, 2010, and 2013. The 2005 increase may be related to the 2000 regularization of several hundred thousand immigrants, while the 2010 increase relates to the 2005 regularization under study here. The 2013 spike in citizenship granting is due to solving technical and bureaucratic issues that had delayed issuance of citizenship for many immigrants.[44]

In Figure 1.5 I show the effect of the 2005 regularization on total drug confiscations, which declined significantly in 2010 and stayed low thereafter. Moreover, this decline came primarily from declines in cocaine confiscations, as shown in Figure 1.6. The decline in cocaine confiscations is consistent with the increase in naturalizations for Latin Americans but a modest decrease in naturalizations for immigrants from Africa, as shown in Figure 1.7.

[44] This is based on a conversation with an employee at Spain's National Statistics Institute.

Overall, these results suggest that granting legal status to immigrants plays an important role in reducing drug trafficking by putting them on a path to citizenship. Taking the average of the coefficients from 2010 to 2016 for the event study estimated on the extensive margin of trafficking suggests that a province granting legal status and subsequent citizenship to an additional 10,000 immigrants reduces the likelihood of a confiscation occurring in that province by 2.3%.

These results differ somewhat from the literature on immigrant legal status on crime. Freedman et al. (2018), Mastrobuoni and Pinotti (2015), and Pinotti (2017) find an immediate drop in immigrant criminal activity as a result of legalization, whereas I find a delayed effect. Pinotti (2017) provides a useful comparison. He shows that for immigrants with weak ties to the formal labor market, legalizations' impact on crime is substantial, but he finds no effect for those with the strongest ties to the formal labor market he finds no effect. Similarly, the 2005 regularization that I study only grants legal status to immigrants with a labor contract already lined up, often a labor relationship that pre-existed 2005 but is simply being formalized by the program. My results are therefore in line with Pinotti (2017) in terms of the immediate effects of legalization, but I look at an extended time horizon and find a reduction in crime around the time immigrants become eligible for citizenship. Therefore the results I present here may be a lower bound on the effects of immigrant legalization on crime.

To get a sense of how the regularization would have affected drug trafficking if it were not conditioned on immigrants' formal labor sector ties, I do a back-of-the-envelope calculation using the instrumented gravity estimates from Section 1.5.2. Using the data and method described in Section 1.5.1, I estimate that in 2004 about a third of all immigrants in Spain were irregular. In addition, using the coefficient on irregular immigrants from column 4 of Table 1.8, I estimate that the regularization program reduced illegal drug trafficking by about 20%.[45] Note, however, that this back-of-the-envelope calculation abstracts away from any offsetting general equilibrium effects.

1.6 Conclusion

The effect of immigration on crime has long been a controversial political issue. In this paper, I contribute to this debate by causally estimating that international immigration is an important factor shaping international drug trafficking, on par with the effect immigrants have on legal trade. This effect is driven primarily by immigrants without legal status, and my evidence shows that granting legal status and a path to citizenship to immigrants can significantly diminish this relationship.

The results presented here have significant relevance to ongoing debates on immigra-

[45] $\exp(5.459 \times (\log(1 + 134 \times \frac{2}{3}/1000) - \log(1 + 134/1000)) - 1 \approx 0.197$.

tion policy in the United States and around the world. In particular, as many European countries and the United States discuss providing some form of amnesty and a path to citizenship to their large populations of undocumented immigrants, this paper offers an additional potential benefit to society from such amnesties. Providing amnesty is also likely to be much cheaper than attempting to keep irregular immigrants from entering the country, such as building a wall. For example, Allen et al. (2018) estimate that the 2007–2010 expansion of the border wall on the U.S.-Mexico border cost approximately $57,500 per deterred immigrant.

An important caveat is that immigrants generate a range of effects on their host countries, from native-born wages to innovation to consumer choice. Hence, generalizing welfare effects of immigration from just one outcome, as is the subject of the present study, may lead to suboptimal policy choices. Instead, policymakers must weigh the varied impacts of migration when shaping immigration policy.

This paper suggests several lines of future research. Subsequent studies in different contexts would be helpful for understanding the external validity of these results. For example, Spain is particularly generous to immigrants in terms of healthcare access relative to many other immigrant-receiving countries, and this may shape the strength of the relationship between legal status and trafficking. In addition, policymakers would benefit from a better understanding of the relative costs and benefits of drug-specific enforcement policies as compared to immigration policies in combating illegal trafficking.

Table 1.1: Effect of Immigrants on Drug Confiscations (OLS)

	Outcome: Confiscations of Imported Drugs 2011-2016 (Dummy)			
	(1)	(2)	(3)	(4)
Log immigrant population	0.0351***	0.0241***	0.0315***	0.00475**
	(0.00342)	(0.00475)	(0.00274)	(0.00157)
Observations	5564	5564	5564	5564
	Confiscations of Drugs Intended for Re-Export (Dummy)			
Log immigrant population	0.0216***	0.0214**	0.0167***	0.00738**
	(0.00422)	(0.00704)	(0.00212)	(0.00217)
Observations	5564	5564	5564	5564
	Confiscations of Imported Drugs 2011-2016 (Log Value)			
Log immigrant population	0.451***	0.313***	0.397***	0.0425
	(0.0534)	(0.0697)	(0.0412)	(0.0212)
Observations	5564	5564	5564	5564
	Confiscations of Drugs Intended for Re-Export 2011-2016 (Log Value)			
Log immigrant population	0.268***	0.258**	0.209***	0.0844***
	(0.0535)	(0.0852)	(0.0301)	(0.0241)
Observations	5564	5564	5564	5564
Country FE	No	Yes	No	Yes
Province FE	No	No	Yes	Yes
Log dist	Yes	Yes	Yes	Yes

Notes: The table presents OLS estimates of equation 1.2 at the country-province level. Standard errors are clustered by 52 provinces in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.2: Effect of Immigrants on Drug Import Confiscations

	(1) First-stage: Log Immigrants 2011	Drug Confiscations 2011-2016 (2) 2SLS: (dummy)	(3) 2SLS: (ln value)
Predicted immigration, 1991-2001	-0.0000641***		
	(0.0000140)		
Predicted immigration, 2001-2011	-0.0000179*		
	(0.0000100)		
Log immigrant population		-0.249*	-5.259**
		(0.135)	(2.245)
Observations	5564	5564	5564
R^2	0.729	-3.175	-9.088
Origin FE	Y	Y	Y
Dest. FE	Y	Y	Y
Ln dist.	Y	Y	Y
1st-stg F-stat.	12.2	12.2	12.2

Notes: The table presents coefficient estimates from IV regressions of equation 1.2 at the country-province level. I instrument for *Log Immigrants 2011* using $\{IV_{o,d}^D = I_{o,-a(d)}^D \times I_{-c(o),d}^D / I_{-c(o)}^D\}_{1991-2001,2001-2011}$ as the excluded instruments, with the first-stage shown in column 1. The dependent variable is a dummy for whether any drugs from country o were confiscated in province d between 2011 and 2016 in column 2 and the log value (in 2012 USD) of drugs from country o confiscated in province d between 2011 and 2016 plus 1. All regressions control for log distance. Standard errors are clustered at the province level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.3: Effect of Immigrants on Drug Re-Export Confiscations

	(1) First-stage: Log Immigrants 2011	Drug Confiscations 2011-2016 (2) 2SLS: (dummy)	Drug Confiscations 2011-2016 (3) 2SLS: (ln value)
Predicted immigration, 1991-2001	-0.0000641***		
	(0.0000140)		
Predicted immigration, 2001-2011	-0.0000179*		
	(0.0000100)		
Log immigrant population		-0.303**	-4.800***
		(0.119)	(1.647)
Observations	5564	5564	5564
R^2	0.729	-6.213	-10.124
Origin FE	Y	Y	Y
Dest. FE	Y	Y	Y
Ln dist.	Y	Y	Y
1st-stg F-stat.	12.2	12.2	12.2

Notes: The table presents coefficient estimates from IV regressions of equation 1.2 at the country-province level. I instrument for *Log Immigrants 2011* using $\{IV^D_{o,d} = I^D_{o,-a(d)} \times I^D_{-c(o),d}/I^D_{-c(o)}\}_{1991-2001,2001-2011}$ as the excluded instruments, with the first-stage shown in column 1. The dependent variable is a dummy for whether any drugs intended for re-export to country o were confiscated in province d between 2011 and 2016 in column 2 and the log value (in 2012 USD) of drugs intended for re-export to country o confiscated in province d between 2011 and 2016 plus 1. All regressions control for log distance. Standard errors are clustered at the province level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.4: Effect of Immigrants on Drug Import Confiscations and Legal Imports (GMM)

	(1) Drug Smuggling	(2) Legal Trade
β_1		
	0.137***	
	(0.021)	
π_1		
	0.006**	
	(0.006)	
β_2		
	2.52***	1.365***
	(0.39)	(0.0998)
π_2		
	0.003**	0.0127*
	(0.003)	(0.00679)
Observations	5564	5136

Standard errors clustered by 52 provinces in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Notes: The sample size in column 2 falls relative to column 1 due to miscoding of certain provinces (specifically Ceuta, Melilla, and the Canary Islands) in the AEAT data on legal trade. I do not estimate β_1 and π_1 for legal trade because virtually all bilateral links engage in some trade.

Table 1.5: Effect of Immigrants on Drug Confiscations: Extensive Margin

	Drug Confiscations 2011-2016 (Dummy)	
	(1)	(2)
Log immigrant population	-0.249*	-0.00622
	(0.135)	(0.0273)
Observations	5564	4051
R^2	-3.175	-0.019
Origin FE	Y	Y
Dest. FE	Y	Y
Ln dist	Y	Y
1st-stg F-stat.	12.2	2.2
Sample	All	< 1000 USD seized

Notes: The table presents coefficient estimates from IV regressions of equation (15) at the country-province level. *Log immigrants 2011* is instrumented with the leave-out push-pull IV from equation (3). In column 2, I subset to the set of country-province pairs for which predicted confiscations (using equation 16) fall below \$1,000. Standard errors are clustered by 52 provinces in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.6: Effect of Immigrants on Illegal Drug Activity (Province Panel)

	(1) First-Stage: Log immigrants	(2) 2SLS: Log value confiscated	(3) 2SLS: Log native-born used drugs last 12 mo.	(4) 2SLS: Log native-born ever used drugs	(5) 2SLS: Log native-born drug trafficking arrests
Ethnic Enclave IV	0.164***				
	(0.0419)				
Log immigrant population		25.72**	2.963	5.767	-7.114
		(12.78)	(4.159)	(5.793)	(12.28)
Observations	728	728	310	312	364
Kleibergen-Paap F-stat	15.4	15.4	1.6	1.7	1.7
Province FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

Notes: The table presents coefficient estimates from IV regressions of equation 1.14 at the province-year level. I instrument for *Log Immigrants* using the excluded instrument defined in equation 1.15, with the first-stage shown in column 1. In column 2, the dependent variable is the log of 1 plus the value of illegal drugs confiscated as measured in the UNODC Individual Seizures Data. The dependent variable of columns 3 and 4 is the log number of native-born Spaniards reporting to the EDADES survey that they used drugs in the last 12 months (column 3) or ever (column 4). The dependent variabe of column 5 is the number of Spanish citizens arrested for illegal drug trafficking. Standard errors are clustered at the autonomous community-by-year level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.7: Effect of Immigrants on Illegal Drug Activity (Province Cross-Section)

	(1) First-stage: Log Immigrants 2011	(2) 2SLS: Log cannabis plants seized
Ethnic Enclave IV	0.488***	
	(0.0229)	
Log immigrant population		1.862***
		(0.362)
Observations	1144	50
R^2	0.551	0.559
1st-stg. F-stat	452.6	452.6
Dep. var. mean (unlogged)	8.8e+04	4003

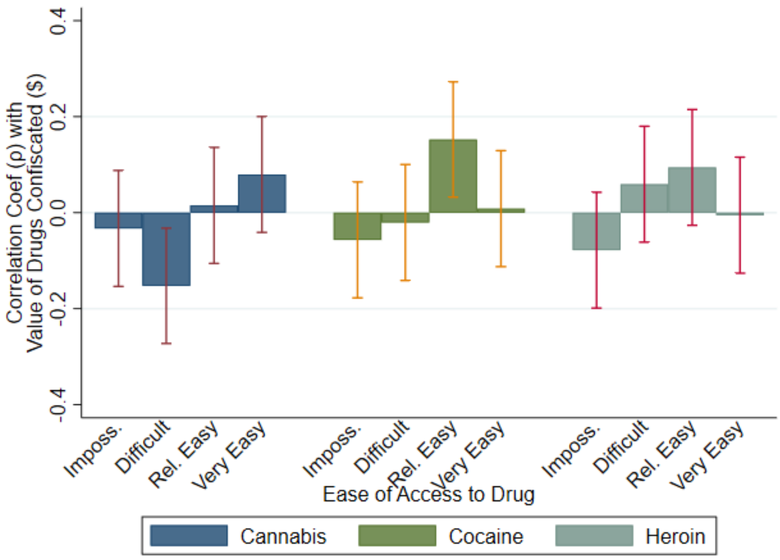
Notes: The table presents coefficient estimates from IV regressions of equation 1.16 at the province level. I instrument for *Log Immigrants 2011* using the excluded instrument defined in equation 1.15, with the first-stage shown in column 1. In column 2, the dependent variable is the log of the number of individuals with Spanish nationality arrested for drug trafficking offenses.Heteroskedasticity-robust standard errors are reported in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.8: Effect of Immigrants by Legal Status on Drug Confiscations

	OLS: Confiscations of Imported Drugs 2011-2016			
	(1)	(2)	(3)	(4)
	(dummy)	(dummy)	(log value)	(log value)
Log immigrant population	0.00432*		0.0322	
	(0.00163)		(0.0218)	
Log regular immigrants		0.0547**		0.526*
		(0.0166)		(0.258)
Log irregular immigrants		0.191***		3.254***
		(0.0263)		(0.307)
Observations	5044	5044	5044	5044
	2SLS: Confiscations of Imported Drugs 2011-2016			
	(dummy)	(dummy)	(log value)	(log value)
Log immigrant population	-0.196		-4.270*	
	(0.103)		(1.698)	
Log regular immigrants		-0.00775		0.344
		(0.0469)		(0.832)
Log irregular immigrants		0.187*		3.444***
		(0.0847)		(0.941)
Observations	5044	5044	5044	5044
Kleibergen-Paap 1st-stg. F-stat.	15.1		15.1	
SW 1st-stg. F-stat. (regular immigrants)		42.0		42.0
SW 1st-stg. F-stat. (irregular immigrants)		168.0		168.0

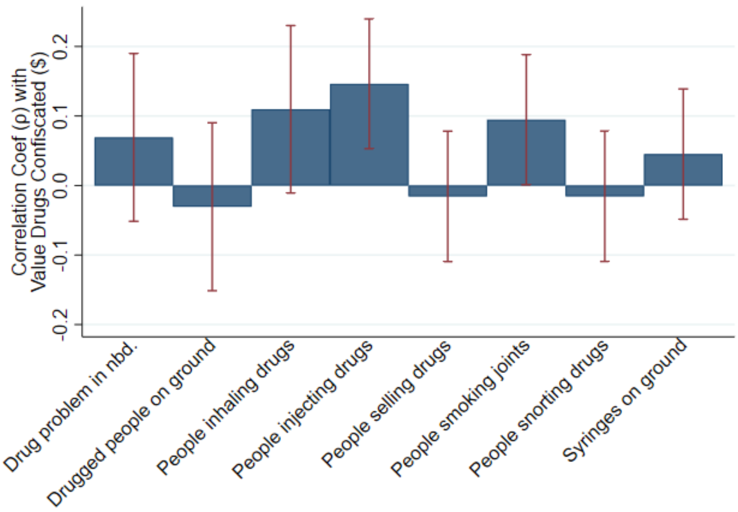
Notes: The table presents estimates of equation 1.18 at the country-province level. Standard errors are clustered by 52 provinces in parentheses. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Figure 1.1: Correlation of Drug Confiscations to Drug Availability by Drug



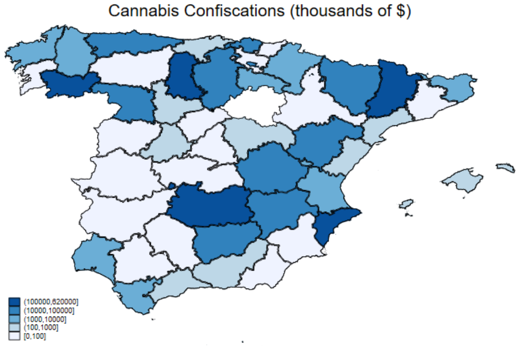
Notes: This figure shows Pearson correlation coefficients between the amount of confiscations per capita of a particular drug with the fraction of respondents in a province who report finding it impossible/difficult/relatively easy/very easy to obtain that drug within 24 hours averaged over the 2011, 2013, and 2015 waves of the EDADES (Suvey on Alcohol and Drugs in Spain) survey. Amphetamines were not asked about until the 2013 survey, and are thus excluded. Ninety percent confidence intervals are shown in red. The sample is a cross-section of 52 Spanish provinces.

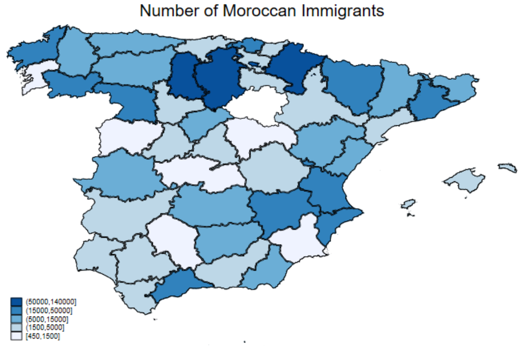
Figure 1.2: Correlation of Drug Confiscations to Drug Availability (across all drugs)



Notes: This figure plots Pearson correlation coefficients between illegal drug confiscations (measured in dollars) per capita across all drugs (as appropriate) with the fraction of respondents in the province who reported observing the listed drug-related behaviors either "frequently" or "very frequently" or, for the first bar, "very." The behaviors listed are, from left to right: (i) "Thinking about where you live, how important of a problem do you think illegal drugs are?"; (ii) "How often in your neighborhood are there drugged people on the ground?"; (iii) "How often in your neighborhood are there people inhaling drugs in paper/aluminium?"; (iv) "How often in your neighborhood are there people injecting drugs?"; (v) "How often in your neighborhood are there people selling drugs?"; (vi) "How often in your neighborhood are there people smoking joints?"; (vii) "How often in your neighborhood are there people snorting drugs by nose?"; (viii) "How often in your neighborhood are there syringes lying on the ground?". As appropriate, I drop cannabis from the drug confiscation variable in the correlation specifically for the questions on people snorting or injecting drugs or syringes being on the ground. Ninety confidence intervals are shown in red. Correlations estimated on a cross-section of 52 Spanish provinces.

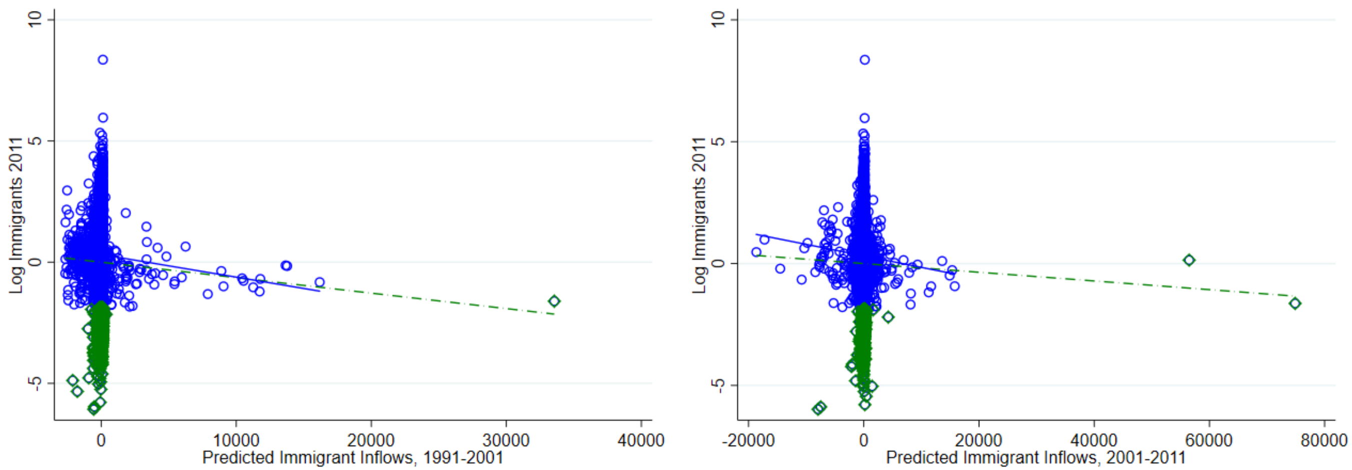
Figure 1.3: Drug Confiscations and Immigrant Population: The Case of Morocco and Cannabis





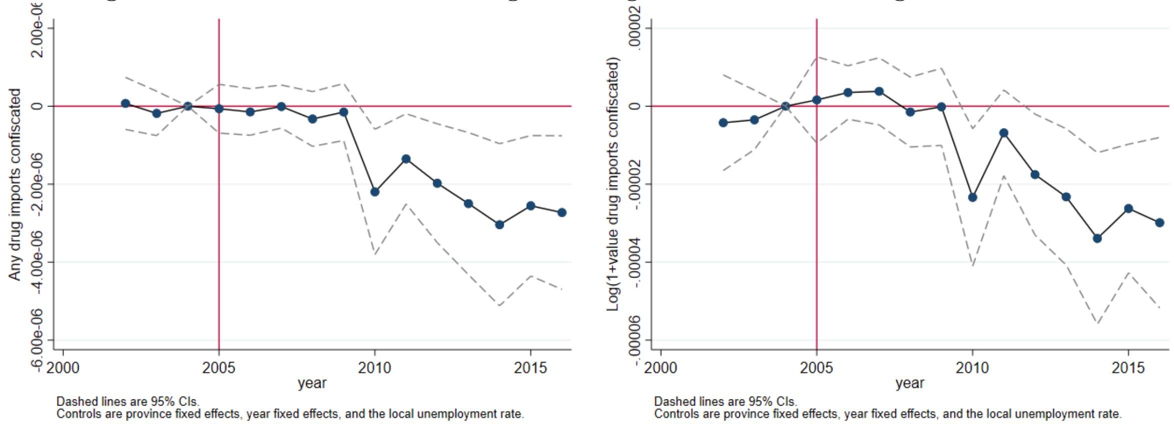
Notes: The figure on the left shows the distribution across Spanish provinces of cannabis confiscations between 2011 and 2016 originating from Morocco; the figure on the right shows the distribution across Spanish provinces of the number of individuals with Moroccan nationality in 2011.

Figure 1.4: First-Stage Fit



Notes: The figure shows the conditional scatter plots of *Log Migrants* 2011 with the instruments for immigrant inflows for decades 1991 to 2001 (on the left) and 2001 to 2011 (on the right). Both *Log Migrants* 2011 and the predicted inflows are residualized on origin and destination fixed effects, log distance, and on the instrument from the left-out decade. I plot the regression line both with (green diamonds, dashed green line) and without (blue circles, blue solid line) outliers.

Figure 1.5: Effect of 2005 Immigrant Regularization on Drug Confiscations

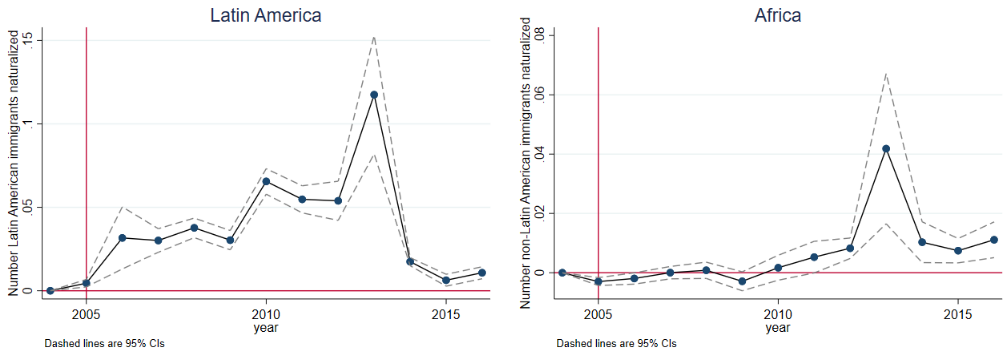


Notes: The figure shows event study plots of the effect of the 2005 immigrant regularization on whether any drugs were confiscated locally (chart on the left) and the log of one plus the value of drugs confiscated locally (chart on the right). Plots are estimated using equation 1.20.

Figure 1.6: Effect of 2005 Immigrant Regularization on Confiscations by Drug Type

Notes: The figure shows event study plots of the effect of the 2005 immigrant regularization on confiscations of cannabis (figures on the left) and cocaine (figures on the right). The dependent variable for the top figures are whether any of the drugs were confiscated locally in that year, and on the bottom the log of one plus the value of the drug confiscated locally. Plots are estimated using equation 1.20.

Figure 1.7: Effect of 2005 Immigrant Regularization on Naturalizations by Continent of Origin



Notes: The figure shows event study plots of the effect of the 2005 immigrant regularization on the number of citizenship acquisitions of immigrants from Latin America (figure on the left) and from Africa (figure on the right). Plots are estimated using equation 1.20.

Appendix

Theory

In this section I briefly lay out a theoretical justification for the bilateral- and province-level regressions discussed above. This theory allows me to provide a structural interpretation to the estimated coefficients from Section 1.3.

Setup. Illegal drug varieties are indexed by $\omega \in [0,1]$ with region d's efficiency in producing variety ω denoted as $z_d(\omega)$. Aggregate consumption of illegal drugs in province d is defined as

$$C_d = \left[\int_0^1 q_d(\omega)^{(\eta-1)/\eta} d\omega\right]^{\eta/(\eta-1)} \tag{1.21}$$

for elasticity of substitution $\eta > 0$ and the quantity of each drug variety $q_d(\omega)$. Following Eaton and Kortum (2002), I assume region d's production efficiency distribution is Frèchet

$$F_d(z) = e^{-T_d z^{-\theta}} \tag{1.22}$$

where $T_d > 0$ and $\theta > 1$ and Z_d has a geometric mean $exp(\gamma/\theta)T_d^{1/\theta}$ where γ is Euler's constant.

In terms of prices, the cost of good ω produced in o and delivered to d is the realization of the random variable

$$P_{od} = \frac{w_o \tau_{od}}{Z_o}$$

for average input wages w_o and bilateral trade costs $\tau_{o,d} \geq 1$ (with $\tau_{dd} = 1$ for all d).

Gravity. Denote by $X_{o,d}$ the flow of illegal drugs from origin country o to destination d. Then I have the gravity equation

$$\ln X_{o,d} = \delta_o + \delta_d + \theta \ln \tau_{o,d}$$

where for bilateral immigrant population $M_{o,d}$,

$$\ln \tau_{o,d} = \alpha_0 \ln t_{o,d} - \alpha_1 \ln M_{o,d} \tag{1.23}$$

where $t_{o,d}$ are bilateral trade costs when the bilateral immigrant population is zero. Hence, we have

$$\ln X_{o,d} = \delta_o + \delta_d + \theta\alpha_0 \ln t_{o,d} - \theta\alpha_1 \ln M_{o,d}$$

In practice, bilateral trade costs (when the bilateral immigrant population is zero) can be expressed as

$$\ln t_{o,d} = f(gravity_{o,d}) + \tilde{\varepsilon}_{o,d}$$

where $f(gravity_{o,d})$ incorporates the standard bilateral gravity variables—geographic or cultural closeness—and $f(\cdot)$ is a standard functional form. Hence, we obtain our estimating equation

$$\ln X_{o,d} = \delta_o + \delta_d + f(gravity_{o,d}) + \beta_2 \ln M_{o,d} + \varepsilon_{o,d} \tag{1.24}$$

where $\varepsilon_{od} \equiv \theta\alpha_0\tilde{\varepsilon}_{o,d}$ and the same applies for $f(\cdot)$ and where $\beta_2 \equiv -\theta\alpha_1$. The unobservable bilateral links that shape trade flows, captured by $\varepsilon_{o,d}$, also shape bilateral migration. Hence, estimating (1.24) using OLS will yield a biased estimate of β_2 (the combination of the trade elasticity and the impact of migration on trade costs). However, with a valid instrument, we can estimate this combination.

Consumption. Following Eaton and Kortum (2002), I have

$$C_d = \frac{1}{\gamma}\left(\frac{T_d}{\pi_{d,d}}\right)^{\frac{1}{\theta}} \tag{1.25}$$

where the share of imports to d coming from o is

$$\pi_{od} = \frac{T_o(w_o\tau_{o,d})^{-\theta}}{\sum_{o'} T_{o'}(w_{o'}\tau_{o',d})^{-\theta}}$$

Assuming $\tau_{d,d} = 1$, I have that

$$\pi_{dd} = \frac{T_d(w_d)^{-\theta}}{\sum_o T_o(w_o\tau_{o,d})^{-\theta}} \tag{1.26}$$

Combining the equations 1.25 and 1.26,

$$C_d = \frac{1}{\gamma} w_d \left(\sum_o T_o(w_o\tau_{o,d})^{-\theta}\right)^{\frac{1}{\theta}}$$

We are interested in understanding the impact of a small change in the vector $\{M_{od}\}_o$ on consumption in d. We assume that $dT_o = 0$ for all $o \neq d$. Log differentiating the previous expression yields

$$d\ln C_d = d\ln w_d + \frac{\pi_{d,d}}{\theta} d\ln T_d - \sum_o \pi_{o,d} d\ln(w_o\tau_{o,d})$$

Now assuming that d is a small economy such that $dw_o = 0$ for all $o \neq d$, we obtain

$$d\ln C_d = (1-\pi_{d,d}) d\ln w_d + \frac{\pi_{d,d}}{\theta} d\ln T_d - \sum_{o\neq d} \pi_{od} d\ln \tau_{o,d}$$

Starting from the previous expression, substituting in equation 1.23 for $d\ln\tau_{o,d}$ to obtain

$$d\ln C_d = (1-\pi_{d,d})d\ln w_d + \frac{\pi_{d,d}}{\theta}d\ln T_d - \sum_{o\neq d}\pi_{od}\left(\alpha_0 d\ln t_{od} - \alpha_1 d\ln M_{o,d}\right)$$

and setting $d\ln t_{od}=0$ (i.e., assuming no change in the impact of time-invariant gravity variables) yields

$$d\ln C_d = (1-\pi_{d,d})d\ln w_d + \frac{\pi_{d,d}}{\theta}d\ln T_d + \alpha_1\sum_{o\neq d}\pi_{o,d}d\ln M_{o,d} + \varepsilon_d$$

where $\varepsilon_d \equiv -\alpha_0\sum_{o\neq d}\pi_{o,d}d\ln\tilde{\varepsilon}_{o,d}$.

To obtain a cross-sectional estimating equation comparable to what I estimate at the province level, I integrate up to obtain

$$\ln C_d - B_0 = (1-\pi_{dd})(\ln w_d + B_1) + \frac{\pi_{d,d}}{\theta}(\ln T_d + B_2) + \alpha_1\sum_{o\neq d}\pi_{o,d}(\ln M_{o,d} + B_o) + \int\varepsilon_d$$

$$\ln C_d = (1-\pi_{d,d})\ln w_d + \frac{\pi_{d,d}}{\theta}\ln T_d + \alpha_1\sum_{o\neq d}\pi_{o,d}\ln M_{o,d}$$

$$+\left(\frac{B_2}{\theta} - B_1\right)\pi_{d,d} + \alpha_1\sum_{o\neq d}B_o\pi_{o,d} + \epsilon_{od}$$

Consider the case of cocaine, where there is no domestic production; that is, $T_d=0$, which implies $\pi_{d,d}=0$. Then we have

$$\ln C_d = \ln w_d + \alpha_1\sum_{o\neq d}\pi_{o,d}\ln M_{o,d} + \alpha_1\sum_{o\neq d}B_o\pi_{o,d} + \tilde{\epsilon}_{od}$$

Finally, to relate consumption as defined in equation 1.25 to empirically observed measures of drug consumption $\tilde{C}_d$, I assume

$$\ln C_d = -\rho_0 + \rho_1\ln\tilde{C}_d$$

Then we have

$$\ln\tilde{C}_d = \rho_0 + \frac{1}{\rho_1}\ln w_d + \frac{\alpha_1}{\rho_1}\sum_{o\neq d}\pi_{o,d}\ln M_{o,d} + \frac{\alpha_1}{\rho_1}\sum_{o\neq d}B_o\pi_{o,d} + \tilde{\epsilon}_{o,d}$$

Additional Empirical Analyses

2004 Madrid Bombing Event Study

I also explore the short-run effects of a major event in Spain: the 2004 Madrid train bombings. Carried out by a Moroccan immigrant and funded by drug trafficking, the bombings killed 193 people, injured about 2,000, and were a major international news story. Due to the connection between the bombings and Moroccan drug trafficking, enforcement intensity directly specifically at Moroccan smuggling may have suddenly increased, while the number of Moroccan immigrants (in the short-run) changed only minimally.

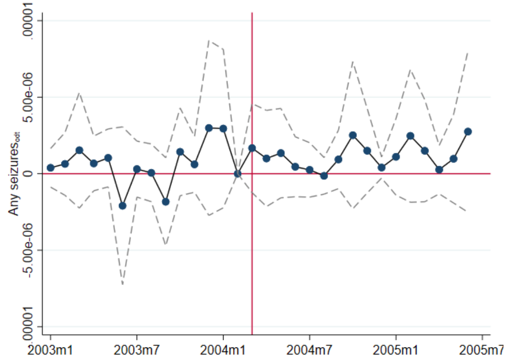
To assess whether this change in enforcement intensity caused a notable increase in drug confiscations, I estimate

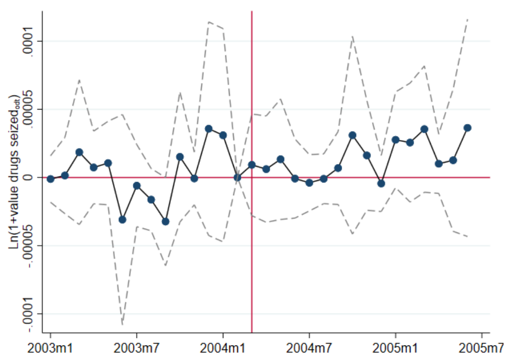
$$Y_{o,d,t} = \alpha_{o,d} + \alpha_t + \sum_{t \neq Mar.\ 2004} \theta_t \times M^{2003}_{Morocco,d} + \epsilon_{o,d,t}$$

where $o \in \{Moroccan, non-Moroccan\}$, d is a Spanish province, t denotes year-month, and $Y_{o,d,t} \in \{\ln(C_{o,d,t,}+1), \mathbf{1}\{C_{o,d,t} > 0\}\}$. The vector $\{\theta_t\}$ will capture the extent to which the number of Moroccan immigrants induces larger changes in enforcement intensity.

I plot the event study graphs in Figure 1.8 and find no statistically significant structural break in confiscations. One caveat for this approach is that the same pattern may result if drug traffickers also suddenly change their trafficking behavior and routes to avoid increased enforcement intensity.

Figure 1.8: Effect of 2005 Bombing on Confiscations from Morocco





Notes: This figure shows event study plots of the effect of the 2004 Madrid train bombings on confiscations of drugs coming from Morocco. I control for year-month and province-by-origin fixed effects, where origins are aggregated into two groups: Moroccan or non-Moroccan. The year-month coefficients plotted are interacted with the number of Moroccan immigrants present in the province in 2003.

Additional Tables and Figures

Table 1.9: Robustness to Different Functional Forms, Any Confiscation

	Drug Confiscations 2011-2016 (Imports, Dummy)		
	(1)	(2)	(3)
$M_{o,d}^{2011}$	0.00000461*		
	(0.00000260)		
$\ln\left(\frac{M_{o,d}^{2011}}{1000}\right)$ (-1 for ∞)		0.0941***	
		(0.0255)	
$\left(M_{o,d}^{2011}\right)^{1/3}$			0.0150**
			(0.00609)
Observations	5564	5564	5564
Country FE	Y	Y	Y
Province FE	Y	Y	Y
Ln dist	Y	Y	Y
1st-stg F-stat.	283.3	7.4	51.9

Standard errors clustered by 52 provinces in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.10: Robustness to Different Functional Forms, Value of Confiscation

	Log(1+Value Confiscated)			Value Confiscated (PPML)
	(1)	(2)	(3)	(4)
$M_{o,d}^{2011}$	0.0000986**			
	(0.0000421)			
$\ln\left(\frac{M_{o,d}^{2011}}{1000}\right)$ (-1 for ∞)		1.967***		
		(0.340)		
$\left(M_{o,d}^{2011}\right)^{1/3}$			0.314***	
			(0.0888)	
Log immigrant population				0.662***
				(0.181)
Observations	5564	5564	5564	3224
Country FE	Y	Y	Y	Y
Province FE	Y	Y	Y	Y
Ln dist	Y	Y	Y	Y
1st-stg F-stat.	283.3	7.4	51.9	

Standard errors clustered by 52 provinces in parentheses. Columns 1–3 are estimated via 2SLS using 1.2 with dependent variable of the log value of one plus illegal drug confiscations. Column 4 estimates 1.2 via Poisson Pseudo-Maximum Likeilihood (PPML) with dependent variable of the value of illegal drug confiscations.

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.11: Effect of Bilateral Immigrant Population by Origin Drug-Hubness

	Drug Confiscations 2011-2016			
	(1)	(2)	(3)	(4)
	Dummy	Log Value	Dummy	Log Value
Log immigrant population	-0.0406	-1.256	0.0193	0.986
	(0.0444)	(0.941)	(0.0757)	(1.210)
Log immigrant population × % of seized drugs from o	-0.0882	-1.287		
	(0.0720)	(2.054)		
Log immigrant population × Drug hubness rank			-0.0000316	-0.0304**
			(0.000531)	(0.0137)
Observations	640	640	640	484
R^2	-0.051	-0.231	0.001	-0.159
Origin FE	Y	Y	Y	Y
Dest. FE	Y	Y	Y	Y
Ln dist	Y	Y	Y	Y
1st-stg F-stat.	2.6	2.6	4.3	5.6

Notes: The table presents coefficient estimates from IV regressions of equation 1.2, modified to include a term interacting the log immigrant population with a measure of the immigrants' origin country drug-hubness at the country-province level. I instrument for *Log Immigrants 2011* using the IV defined in equation 1.4 and the IV interacted with the measure of drug hubness. The dependent variable is a dummy for whether any illegal drugs imported from country o were confiscated in province d between 2011 and 2016 in columns 1 and 3, and the log of 1 plus the value (in 2012 USD) of illegal drugs imported from country o were confiscated in province d between 2011 and 2016 in columns 2 and 4. All regressions control for nationality and province fixed effects as well as log distance. Standard errors are clustered at the province level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.12: Gravity Specification: Alternative Standard Errors

		Drug Confiscations, 2011-2016			
	(1)	(2)	(3)	(4)	(5)
	First-stage: Log Immigrants 2011	2SLS: dummy imports	2SLS: log value imports	2SLS: dummy re-export	2SLS: log value re-export
PANEL A: HETEROSKEDASTICITY-ROBUST					
Predicted immigration, 1991-2001	-0.0000641***				
	(0.0000116)				
Predicted immigration, 2001-2011	-0.0000179*				
	(0.00000799)				
Log immigrant population		-0.249*	-5.259*	-0.303*	-4.800**
		(0.121)	(2.106)	(0.122)	(1.668)
Constant	3.247**				
	(1.204)				
Kleibergen-Paap F-stat.	10.4	15.6	15.6	15.6	15.6
PANEL B: CLUSTERED BY COUNTRY					
Predicted immigration, 1991-2001	-0.0000641***				
	(0.0000130)				
Predicted immigration, 2001-2011	-0.0000179***				
	(0.00000372)				
Log immigrant population		-0.249	-5.259	-0.303**	-4.800*
		(0.206)	(3.335)	(0.107)	(1.841)
Constant	3.247**				
	(0.988)				
Kleibergen-Paap F-stat.	13.0	19.2	19.2	19.2	19.2
PANEL C: CLUSTERED BY PROVINCE (BASELINE)					
Predicted immigration, 1991-2001	-0.0000641***				
	(0.0000140)				
Predicted immigration, 2001-2011	-0.0000179				
	(0.0000100)				
Log immigrant population		-0.249	-5.259*	-0.303*	-4.800**
		(0.135)	(2.245)	(0.119)	(1.647)
Constant	3.247				
	(1.813)				
Kleibergen-Paap F-stat.	8.3	12.2	12.2	12.2	12.2
PANEL D: CLUSTERED TWO-WAYS BY COUNTRY AND PROVINCE					
Predicted immigration, 1991-2001	-0.0000641***				
	(0.0000151)				
Predicted immigration, 2001-2011	-0.0000179*				
	(0.00000699)				
Log immigrant population		-0.249	-5.259	-0.303**	-4.800*
		(0.214)	(3.418)	(0.104)	(1.816)
Constant	3.247				
	(1.663)				
Kleibergen-Paap F-stat.	6.9	9.8	9.8	9.8	9.8

Notes: The table presents regression results at the country-province level for the first-stage (column 1) and the second stage (columns 2–5) of the baseline gravity specification. All regressions control for province and nationality fixed effects and log distance. In Panel A, I compute heteroskedasticity-robust standard errors with no clustering. In Panel B, I cluster by nationality; in Panel C by province, as in the baseline specification; and in Panel D, I cluster two-ways by country and province. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 1.13: Effect of Immigrants on Import Confiscations: Panel Analysis (no o, d fixed effects)

		Import Drug Confiscations 2011-2016			
	(1) First-stage: Log Immigrants	(2) OLS: dummy	(3) OLS: log value	(4) 2SLS: dummy	(5) 2SLS: log value
Predicted immigration, 1991-2001	0.000134***				
	(0.0000279)				
Predicted immigration, 2001 to t	0.0000796***				
	(0.0000112)				
(Predicted immigration, 1991-2001)^2	-5.35e-09***				
	(9.63e-10)				
(Predicted immigration, 2001 to t)^2	-4.50e-10***				
	(6.82e-11)				
Log immigrant population		0.101***	1.385***	0.188***	2.636***
		(0.00925)	(0.142)	(0.0218)	(0.313)
Observations	80080	80080	80080	80080	80080
Origin-Year FE	Y	Y	Y	Y	Y
Dest.-Year FE	Y	Y	Y	Y	Y
Origin-Dest. FE	N	N	N	N	N
1st-stg F-stat.	60.1			48.8	48.8

Standard errors clustered two-ways at the year-province and origin-destination levels in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.14: Effect of Immigrants on Import Confiscations: Panel Analysis (with o, d fixed effects)

		Import Drug Confiscations 2011-2016			
	(1) First-stage: Log Immigrants	(2) OLS: dummy	(3) OLS: log value	(4) 2SLS: dummy	(5) 2SLS: log value
Predicted immigration, 2001 to t	0.0000131***				
	(0.00000230)				
(Predicted immigration, 2001 to t)^2	-6.68e-11***				
	(1.26e-11)				
Log immigrant population		0.0477***	1.382***	1.185***	14.25***
		(0.0147)	(0.141)	(0.367)	(4.493)
Observations	80080	85800	85800	80080	80080
Origin-Year FE	Y	Y	Y	Y	Y
Dest.-Year FE	Y	Y	Y	Y	Y
Origin-Dest. FE	Y	Y	Y	Y	Y
1st-stg F-stat.	32.4			16.5	16.5

Standard errors clustered two-ways at the year-province and origin-destination levels in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.15: Effect of Immigrants on Re-Export Confiscations: Panel Analysis (no o, d fixed effects)

	(1) First-stage: Log Immigrants	Re-Export Drug Confiscations 2011-2016 (2) OLS: dummy	(3) OLS: log value	(4) 2SLS: dummy	(5) 2SLS: log value
Predicted immigration, 1991-2001	0.000134***				
	(0.0000279)				
Predicted immigration, 2001 to t	0.0000796***				
	(0.0000112)				
(Predicted immigration, 1991-2001)^2	-5.35e-09***				
	(9.63e-10)				
(Predicted immigration, 2001 to t)^2	-4.50e-10***				
	(6.82e-11)				
Log immigrant population		0.0257***	0.311***	0.0525***	0.638***
		(0.00494)	(0.0605)	(0.0153)	(0.188)
Observations	80080	85800	85800	80080	80080
Origin-Year FE	Y	Y	Y	Y	Y
Dest.-Year FE	Y	Y	Y	Y	Y
Origin-Dest. FE	N	N	N	N	N
1st-stg F-stat.	60.1			48.8	48.8

Standard errors clustered two-ways at the year-province and origin-destination levels in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.16: Effect of Immigrants on Re-Export Confiscations: Panel Analysis (with o, d fixed effects)

		Re-Export Drug Confiscations 2011-2016			
	(1) First-stage: Log Immigrants	(2) OLS: dummy	(3) OLS: log value	(4) 2SLS: dummy	(5) 2SLS: log value
Predicted immigration, 2001 to t	0.0000131***				
	(0.00000230)				
(Predicted immigration, 2001 to t)^2	-6.68e-11***				
	(1.26e-11)				
Log immigrant population		-0.0137	0.310***	-0.168	-1.914
		(0.00881)	(0.0606)	(0.171)	(2.047)
Observations	80080	85800	85800	80080	80080
Origin-Year FE	Y	Y	Y	Y	Y
Dest.-Year FE	Y	Y	Y	Y	Y
Origin-Dest. FE	Y	Y	Y	Y	Y
1st-stg F-stat.	32.4			16.5	16.5

Standard errors clustered two-ways at the year-province and origin-destination levels in parentheses

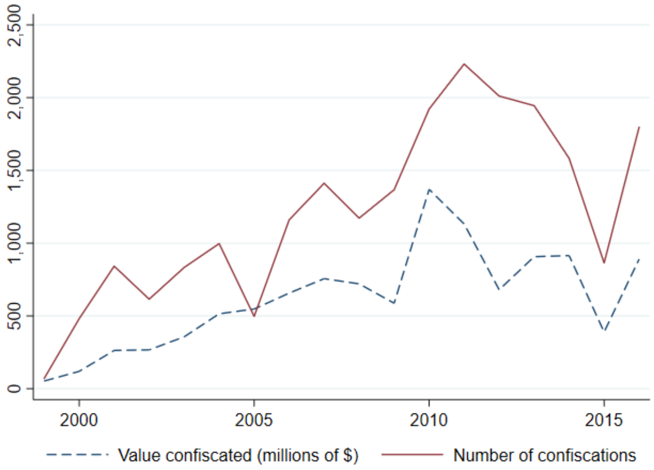
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 1.17: Effect of Immigrants on Illegal Drug Activity: Province Panel with Leave-Out Instrument

	(1) First-Stage: Log immigrants	(2) 2SLS: Log value confiscated	(3) 2SLS: Log native-born used drugs last 12 mo.	(4) 2SLS: Log native-born ever used drugs	(5) 2SLS: Log native-born drug trafficking arrests
Shift-Share IV	0.135^{***}				
	(0.0427)				
Log immigrant population		29.21^{*}	4.286	7.901	-15.92
		(16.07)	(6.722)	(9.876)	(35.30)
Observations	728	728	310	312	364
Kleibergen-Paap F-stat	10.0	10.0	0.9	0.9	0.5
Province FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes

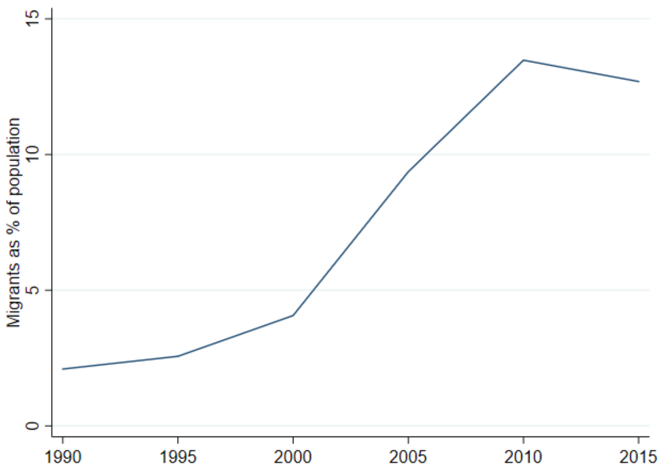
Notes: The table presents coefficient estimates from IV regressions of equation 1.14 at the province-year level. I instrument for *Log Immigrants* using the excluded instrument defined in equation 1.15, with the first-stage shown in column 1. In column 2, the dependent variable is the log of 1 plus the value of illegal drugs confiscated as measured in the UNODC Individual Seizures Data. The dependent variable of columns 3 and 4 is the log number of native-born Spaniards reporting to the EDADES survey that they used drugs in the last 12 months (column 3) or ever (column 4). The dependent variabe of column 5 is the number of Spanish citizens arrested for illegal drug trafficking. Standard errors are clustered at the autonomous community-by-year level. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Figure 1.9: Illegal Drug Confiscations per Year, 1999-2016



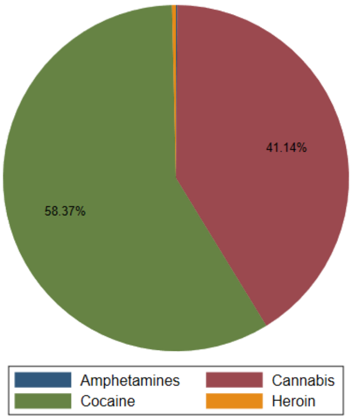
Notes: This figure shows the value of drugs trafficked from foreign countries confiscated over time by Spanish authorities and the number of confiscation events as reported to the United Nations Office of Drugs and Crime (UNODC). Drug prices used are 2012 wholesale prices taken from a survey of Spanish drug prices reported to the UNODC.

Figure 1.10: Immigrant Population Share in Spain, 1990–2015



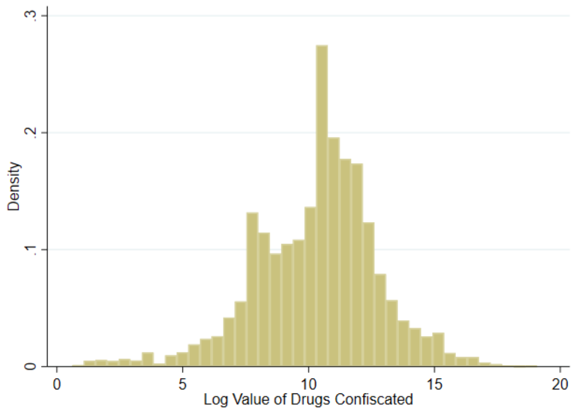
Notes: This figure shows the fraction of the Spanish population born in another country over time. The data are reported by the World Bank but originally come from the United Nations Population Division.

Figure 1.11: Confiscations by Drug Type



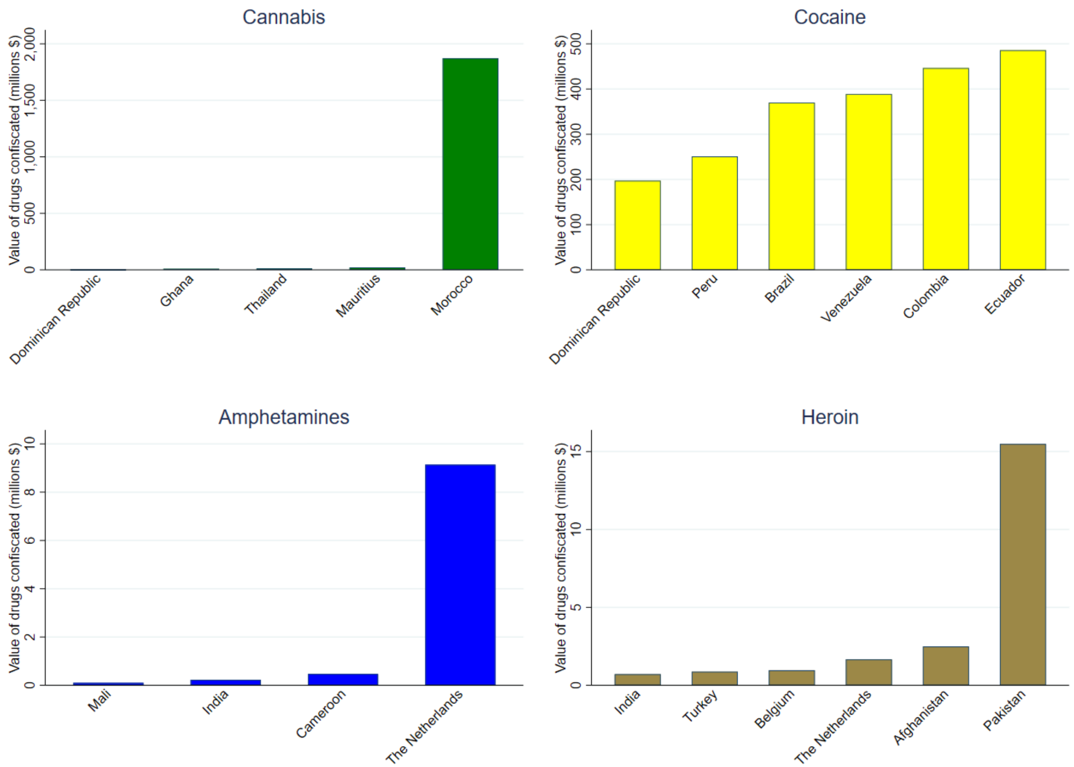
Notes: This figure shows the makeup of drug confiscations in Spain by drug type. Drug prices used are 2012 wholesale prices taken from a survey of Spanish drug prices reported to the United Nations Office of Drugs and Crime (UNODC).

Figure 1.12: Distribution of Log Value of Confiscations



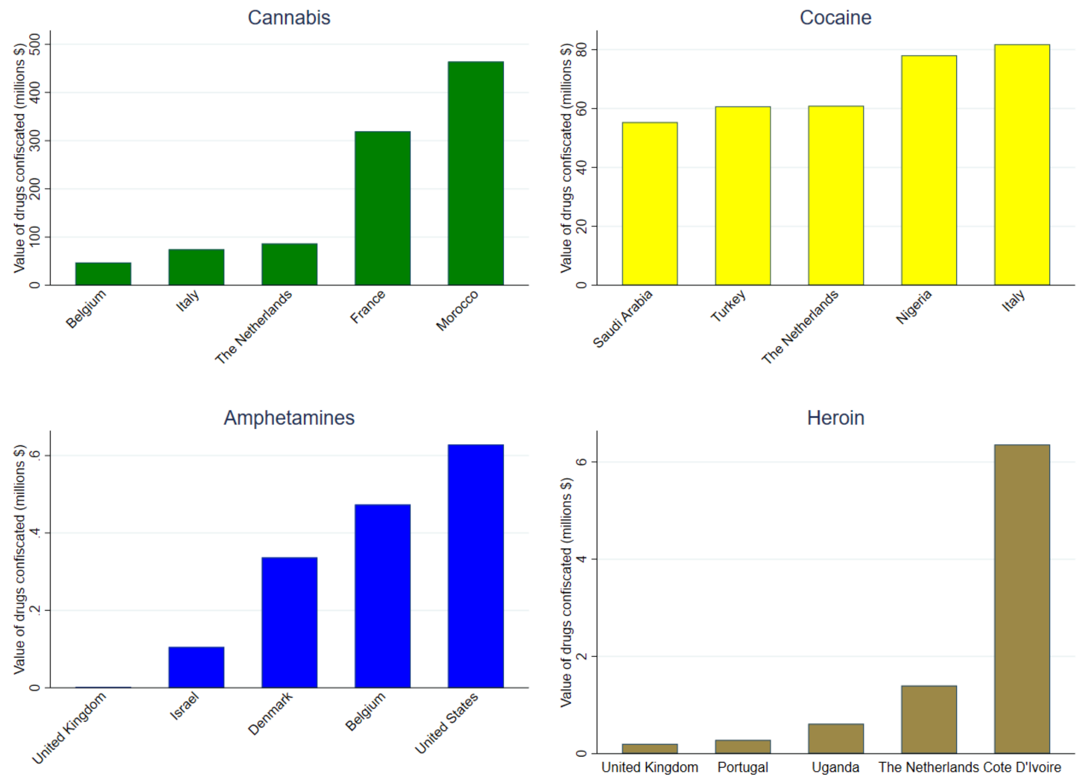
Notes: This figure shows the distribution of the log value of drug confiscations in Spain between 2011 and 2016 as reported to the United Nations Office of Drugs and Crime (UNODC). Drug prices used are 2012 wholesale prices taken from a survey of Spanish drug prices reported to the UNODC.

Figure 1.13: Top Five Origins by Drug



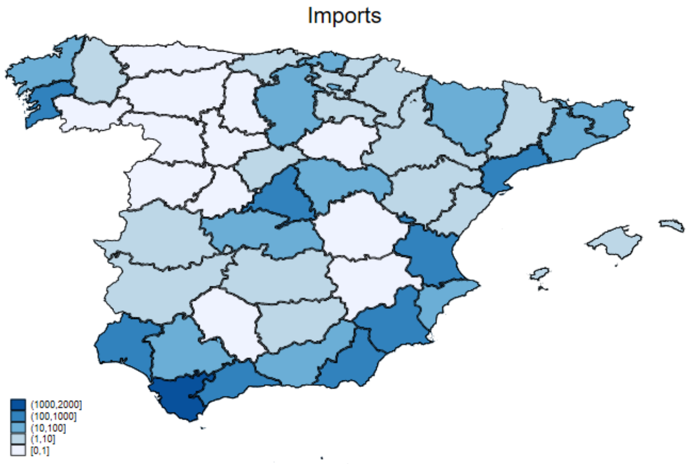
Notes: This figure shows the top five exporters of illegal drugs to Spain during 2011 through 2016 by drug as proxied by confiscations by Spanish law enforcement reported to the United Nations Office of Drugs and Crime (UNODC).

Figure 1.14: Top 5 Intended Destinations by Drug



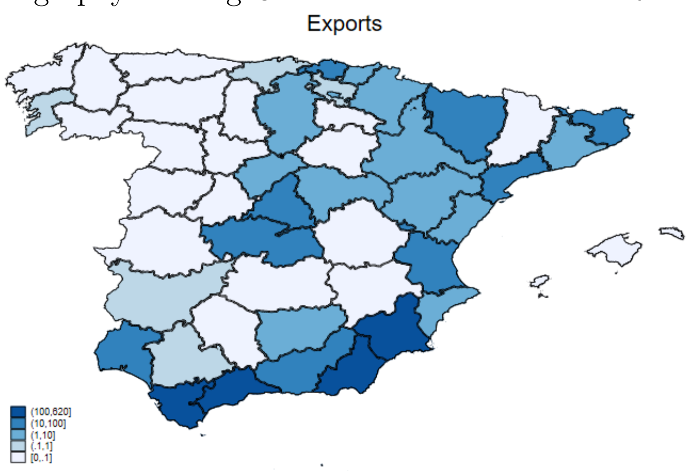
Notes: This figures shows the top 5 importers of illegal drugs from Spain during 2011 through 2016 by drug as proxied by confiscations by Spanish law enforcement reported to the United Nations Office of Drugs and Crime.

Figure 1.15: Geography of Drug Import Confiscations in Spain



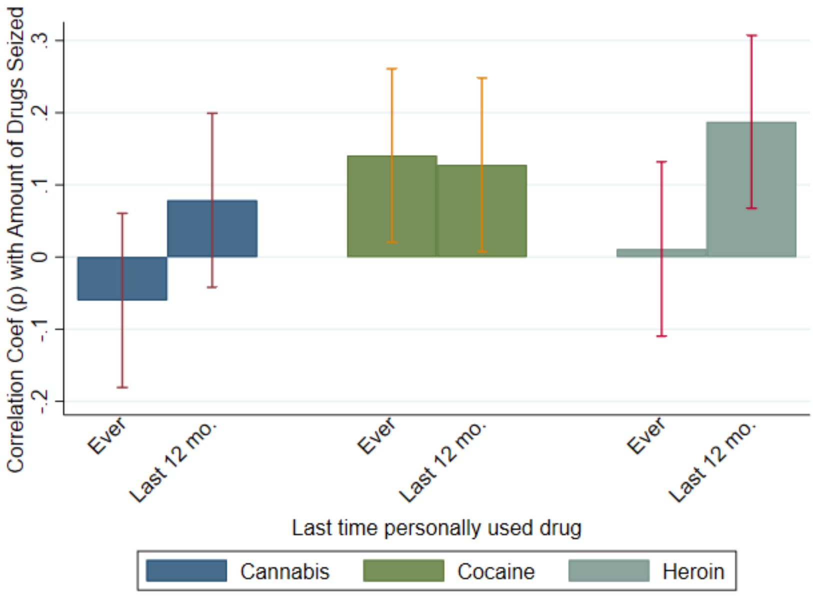
Notes: This figure shows the distribution of drug confiscations of imports (measured in dollars by the estimated wholesale value of confiscated drugs) per capita across Spanish provinces for confiscations occurring between 2011 and 2016 as reported by Spain to the United Nations Office of Drugs and Crime.

Figure 1.16: Geography of Drug Confiscations Intended for Re-Export in Spain



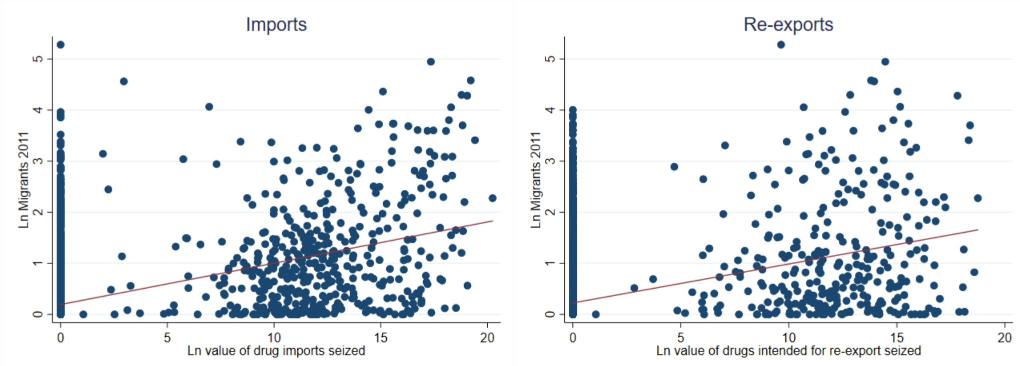
Notes: This figure shows the distribution of confiscations of drugs intended for re-export (measured in dollars by the estimated wholesale value of confiscated drugs) per capita across Spanish provinces for confiscations occurring between 2011 and 2016 as reported by Spain to the United Nations Office of Drugs and Crime.

Figure 1.17: Correlation of Drug Confiscations to Personal Use by Drug



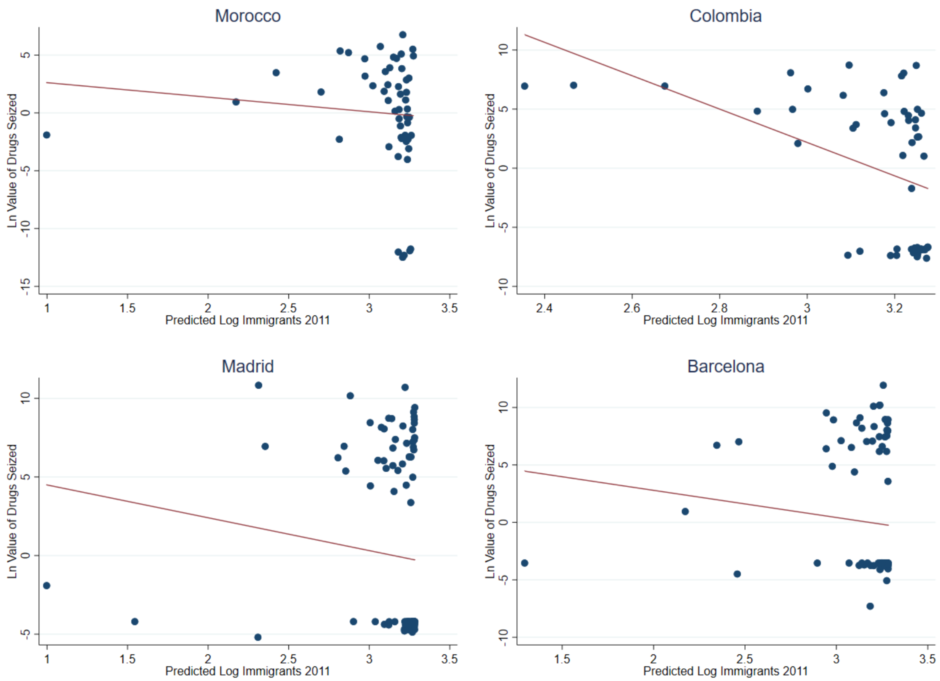
Notes: This figure shows the correlation coefficient between the amount confiscated per capita of a particular drug with the fraction of respondents in a province who report having ever used the drug or having used the drug within the last 12 months averaged over the 2011, 2013, and 2015 waves of the EDADES (Survey on Alcohol and Drugs in Spain) survey. Amphetamines were not asked about until the 2013 survey and are thus excluded. Ninety percent confidence intervals are shown in red. The sample is a cross-section of 52 Spanish provinces.

Figure 1.18: Drug Confiscations and Number of Immigrants Raw Correlation



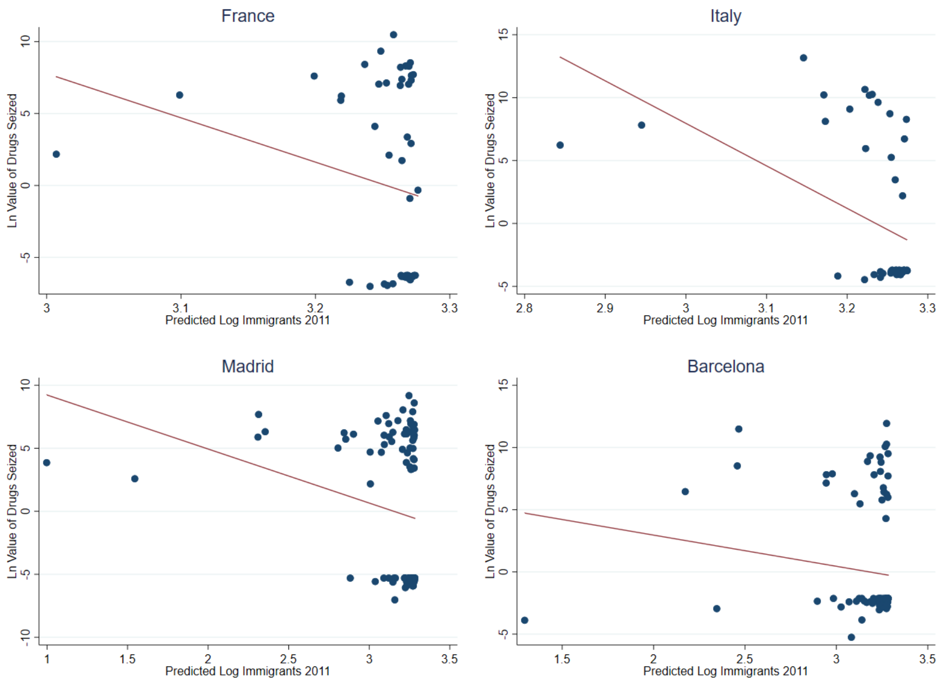
Notes: The figure on the left shows the unconditional scatter plot of the bilateral log value of drug imports confiscated on the x-axis with the bilateral log number of immigrants measured in 2011 on the y-axis. The figure on the right is the same but plots the log of one plus the value of drugs confiscated intended for re-export on the x-axis.

Figure 1.19: Migrants and Drug Trafficking Imports



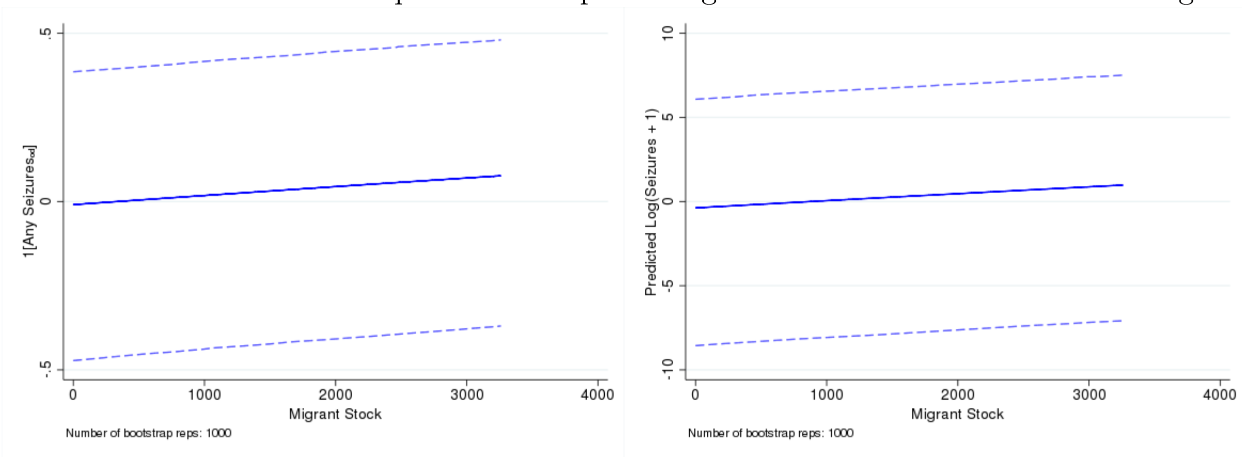
Notes: The figure shows the conditional scatter plots of predicted *Log Migrants* 2011 with the log value of imported drugs confiscated for origins Morocco and Colombia and separately for provinces Madrid and Barcelona. Data are conditional on origin and destination fixed effects and log distance.

Figure 1.20: Migrants and Drug Trafficking Exports



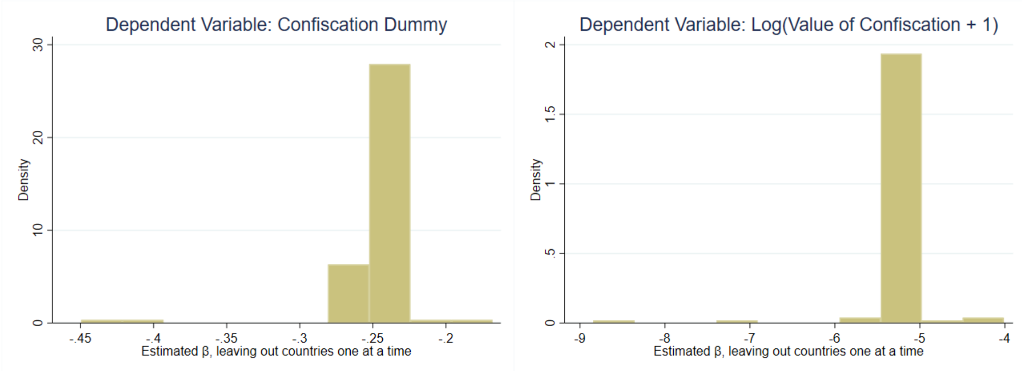
Notes: The figure shows the conditional scatter plots of predicted *Log Migrants* 2011 with the log value of imported drugs confiscated for origins France and Italy and separately for provinces Madrid and Barcelona. Data are conditional on origin and destination fixed effects and log distance.

Figure 1.21: Non-Parametric Relationship between Import Drug Confiscations and Bilateral Immigrant Population



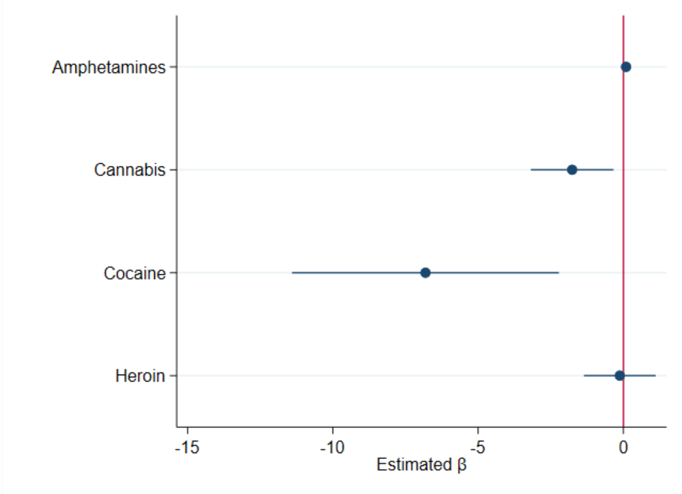
Notes: This figure shows the values of the dummy variable $\mathbf{1}\{S_{od} > 0\}$ (left) or $\log(S_{od} + 1)$ (right) predicted from the non-parametrically estimated function $g(M_{o,d})$, as in $f(S_{o,d}) = \alpha_o + \alpha_d + g(M_{o,d}) + \delta \ln(Dist_{o,d}) + \varepsilon_{o,d}$. S_{od} is equal to the value of drugs confiscated in province d originating from country o. For estimation I used the Stata program `npiv` developed by Chetverikov et al. (2018).

Figure 1.22: Effect of Immigrants on Drug Trafficking: Dropping Origin Countries



Notes: The figures show the distribution of the estimated effect of immigrants on illegal drug confiscations (β from equation 1.2) when leaving out one nationality at a time. The figure on the left shows the distribution of βs when the dependent variable of equation 1.2 is a dummy for whether any drug import from a given origin country was confiscated locally between 2011 and 2016. The figure on the right shows the distribution of β's when the dependent variable is the log of one plus the value of drugs imported from a given origin country and confiscated locally between 2011 and 2016.

Figure 1.23: Effect of Immigrants on Drug Trafficking by Drug



Notes: The figure shows the effect of immigrants on drug trafficking (β from equation 1.2) for the four drugs included for the baseline estimation. As shown in Figure 1.11, cannabis and cocaine make up the vast majority of illegal drugs confiscated by Spanish authorities.

Figure 1.24: Binscatter, Any Confiscation on Bilateral Immigrant Population

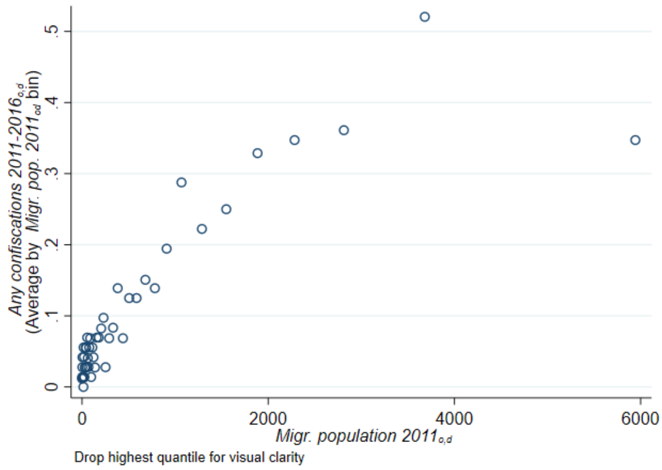
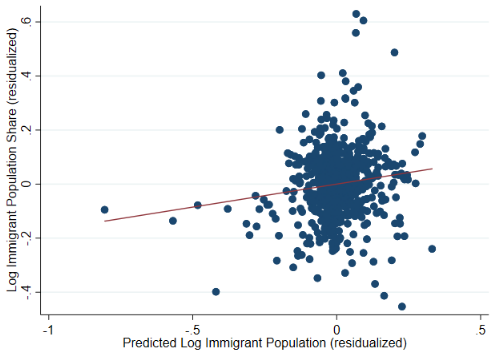
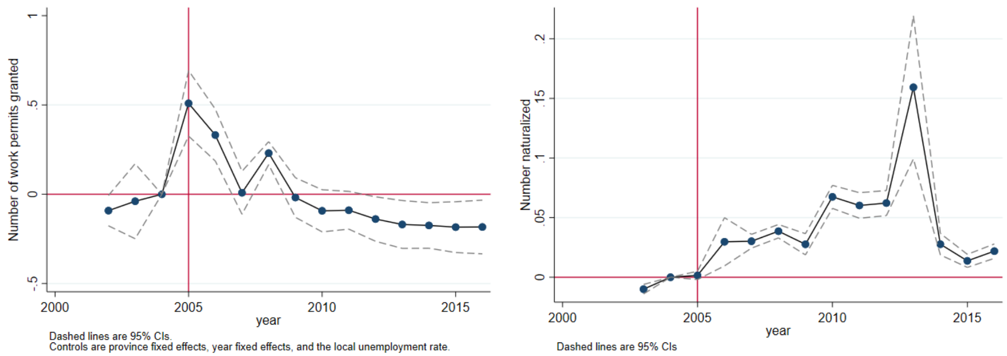


Figure 1.25: First-Stage Fit, Province-Level Panel



Note: The figure shows the first-stage fit of province immigrant population on the province-level shift-share instrumental variable defined in equation 1.15, both residualized on year and province fixed effects.

Figure 1.26: Effect of 2005 Immigrant Regularization on Work Permits, Naturalizations



Notes: The figure shows event study plots of the effect of the 2005 immigrant regularization on the number of residency permits granted to immigrants (chart on the left) and the number of immigrants obtaining citizenship (chart on the right). Plots are estimated using equation 1.20. The large spike in citizenship acquisitions in 2013 was caused by a concerted effort by the Spanish government to reduce delays in processing citizenship applications.

Literature Cited

Abadie, A., M. C. Acevedo, M. Kugler, and J. Vargas (2014). Inside the war on drugs: Effectiveness and unintended consequences of a large illicit crops eradication program in Colombia. Unpublished manuscript.

Akee, R., A. K. Basu, A. Bedi, and N. H. Chau (2014, May). Transnational trafficking, law enforcement, and victim protection: A middleman trafficker's perspective. *Journal of Law and Economics 57*, 349–386.

Allen, T., C. Dobbin, and M. Morten (2018). Border walls. No. w25267, National Bureau of Economic Research.

Alvarez, A., J. F. Gamella, and I. Parra (2016). Cannabis cultivation in Spain: A profile of plantations, growers and production systems. *International Journal of Drug Policy 37*, 70–81.

Atkin, D. (2013). Trade, tastes, and nutrition in India. *American Economic Review 103*(5), 1629–63.

Aziani, A., G. Berlusconi, and L. Giommoni (2019). A quantitative application of enterprise and social embeddedness theories to the transnational trafficking of cocaine in Europe. *Deviant Behavior*, 1–23.

Becker, G. S. (1968). Crime and punishment: An economic approach. In *The Economic Dimensions of Crime*, pp. 13–68. Springer.

Bell, B., F. Fasani, and S. Machin (2013). Crime and immigration: Evidence from large immigrant waves. *Review of Economics and Statistics 21*(3), 1278–1290.

Berlusconi, G., A. Aziani, and L. Giommoni (2017). The determinants of heroin flows in Europe: A latent space approach. *Social Networks 51*, 104–117.

Bhagwati, J. and B. Hansen (1973). A theoretical analysis of smuggling. *Quarterly Journal of Economics 87*, 172–187.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The quarterly journal of economics 118*(4), 1335–1374.

Bronnenberg, B. J., J.-P. H. Dubé, and M. Gentzkow (2012). The evolution of brand preferences: Evidence from consumer migration. *American Economic Review 102*(6), 2472–2508.

Burchardi, K. B., T. Chaney, and T. A. Hassan (2019). Migrants, ancestors, and foreign investments. *The Review of Economic Studies 86*(4), 1448–1486.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics 19*(1), 22–64.

Castillo, J. C., D. Mejía, and P. Restrepo (2020). Scarcity without Leviathan: The violent effects of cocaine supply shortages in the Mexican drug war. *Review of Economics and Statistics 102*(2), 269–286.

Chetverikov, D., D. Kim, and D. Wilhelm (2018). Nonparametric instrumental-variable estimation. *The Stata Journal 18*(4), 937–950.

Chetverikov, D. and D. Wilhelm (2017). Nonparametric instrumental variable estimation under monotonicity. *Econometrica 85*(4), 1303–1320.

Cohen, L., U. G. Gurun, and C. Malloy (2017). Resident networks and corporate connections: Evidence from World War II internment camps. *The Journal of Finance 72*(1), 207–248.

Combes, P.-P., M. Lafourcade, and T. Mayer (2005). The trade-creating effects of business and social networks: evidence from France. *Journal of international Economics 66*(1), 1–29.

Cortes, P. (2008). The effect of low-skilled immigration on US prices: evidence from CPI data. *Journal of Political Economy 116*(3), 381–422.

Dell, M. (2015). Trafficking networks and the Mexican drug war. *American Economic Review 105*(6), 1738–1779.

Díez Nicolás, J. and M. J. Ramírez Lafita (2001). La voz de los inmigrantes. *IMSERSO, Ministerio de Trabajo y Asuntos Sociales*.

Dobkin, C. and N. Nicosia (2009). The war on drugs: Methamphetamine, public health, and crime. *American Economic Review 99*(1), 324–49.

Dube, O., O. Garcia-Ponce, and K. Thom (2016). From maize to haze: Agricultural shocks and the growth of the Mexican drug sector. *Journal of the European Economic Association 14*(5), 1181–1224.

Dunlevy, J. A. (2006). The influence of corruption and language on the protrade effect of immigrants: Evidence from the American states. *Review of Economics and Statistics 88*(1), 182–186.

Dustmann, C., T. Frattini, and I. P. Preston (2013). The effect of immigration along the distribution of wages. *Review of Economic Studies 80*(1), 145–173.

Dustmann, C., U. Schönberg, and J. Stuhler (2016). The impact of immigration: Why do studies reach such different results? *Journal of Economic Perspectives 30*(4), 31–56.

Eaton, J. and S. Kortum (2002). Technology, geography, and trade. *Econometrica 70*(5), 1741–1779.

Ehrlich, I. (1973). Participation in illegitimate activities: A theoretical and empirical investigation. *Journal of Political Economy 81*(3), 521–565.

Esposito, A. and F. Iturrieta (2017, January 25). Chile's presidential hopefuls bet on anti-immigrant sentiment. *Reuters*.

Felbermayr, G., V. Grossmann, and W. Kohler (2015). Migration, international trade, and capital formation: Cause or effect? In *Handbook of the Economics of International Migration*, Volume 1, pp. 913–1025. Elsevier.

Fisman, R. and S.-J. Wei (2009). The smuggling of art, and the art of smuggling: Uncovering the illicit trade in culteral property and antiques. *American Economic Journal: Applied Economics 1*, 82–96.

Freedman, M., E. Owens, and S. Bohn (2018). Immigration, employment opportunities, and criminal behavior. *American Economic Journal: Economic Policy 10*(2), 117–51.

Friedberg, R. M. (2001). The impact of mass migration on the israeli labor market. *The Quarterly Journal of Economics 116*(4), 1373–1408.

Gálvez Iniesta, I. (2020). The size, socio-economic composition and fiscal implications of the irregular immigration in Spain. Unpublished manuscript.

Giommoni, L., A. Aziani, and G. Berlusconi (2017). How do illicit drugs move across countries? a network analysis of the heroin supply to Europe. *Journal of Drug Issues 47*(2), 217–240.

González-Enríquez, C. (2009). Undocumented migration: Counting the uncountable. data and trends across europe; country report: Spain. Technical report, report prepared for the research project CLANDESTINO.

Gould, D. M. (1994). Immigrant links to the home country: Empirical implications for US bilateral trade flows. *The Review of Economics and Statistics*, 302–316.

Grossman, G. M. and C. Shapiro (1988). Foreign counterfeiting of status goods. *Quarterly Journal of Economics 103*(1), 79–100.

Head, K. and T. Mayer (2014). Gravity equations: Workhorse, toolkit, and cookbook. In *Handbook of international economics*, Volume 4, pp. 131–195. Elsevier.

Head, K. and J. Ries (1998). Immigration and trade creation: Econometric evidence from Canada. *Canadian Journal of Economics*, 47–62.

Kaushal, N. (2006). Amnesty programs and the labor market outcomes of undocumented workers. *Journal of Human Resources 41*(3), 631–647.

Kossoudji, S. A. and D. A. Cobb-Clark (2002). Coming out of the shadows: Learning about legal status and wages from the legalized population. *Journal of Labor Economics 20*(3), 598–628.

Marsh, K., L. Wilson, and R. Kenehan (2012). The impact of globalization on the UK market for illicit drugs: Evidence from interviews with convicted drug traffickers. In C. C. Storti and P. D. Grauwe (Eds.), *Illicit Trade and the Global Economy*, CESifo Seminar Series, pp. 159–177. The MIT Press.

Mastrobuoni, G. and P. Pinotti (2015). Legal status and the criminal activity of immigrants. *American Economic Journal: Applied Economics 7*(2), 175–206.

Matrix Knowledge Group (2007). The illicit drug trade in the United Kingdom. Technical report, London: Home Office.

Mavrellis, C. (2017). Transnational crime and the developing world. Technical report, Global Financial Integrity.

Mejía, D., P. Restrepo, and S. V. Rozo (2017). On the effects of enforcement on illegal markets: Evidence from a quasi-experiment in Colombia. *The World Bank Economic Review 31*(2), 570–594.

Minnesota Population Center (2019). Integrated public use microdata series, international. https://doi.org/10.18128/D020.V7.2.

Monras, J. (2020). Immigration and wage dynamics: Evidence from the mexican peso crisis. *Journal of Political Economy 128*(8), 3017–3089.

NDIC (2011, April). The economic impact of illicit drug use on American society. Technical report, U.S. Department of Justice.

OECD (2010). International migration outlook: Sopemi.

Pajares, M. (2004). Inmigración irregular en cataluña. Análisis y propuestas. *Barcelona: CERES*.

Parsons, C. and P.-L. Vézina (2018). Migrant networks and trade: The Vietnamese boat people as a natural experiment. *The Economic Journal 128*(612), F210–F234.

Peri, G. and F. Requena-Silvente (2010). The trade creation effect of immigrants: Evidence from the remarkable case of Spain. *Canadian Journal of Economics 43*(4), 1433–1459.

Pinotti, P. (2017). Clicking on heaven's door: The effect of immigrant legalization on crime. *American Economic Review 107*(1), 138–68.

Rauch, J. E. and V. Trindade (2002). Ethnic Chinese networks in international trade. *Review of Economics and Statistics 84*(1), 116–130.

Sanromá, E., R. Ramos, and H. Simón (2015). How relevant is the origin of human capital for immigrant wages? evidence from Spain. *Journal of Applied Economics 18*(1), 149–172.

Silva, J. S. and S. Tenreyro (2006). The log of gravity. *The Review of Economics and Statistics 88*(4), 641–658.

Silva, J. S. and S. Tenreyro (2011). Further simulation evidence on the performance of the poisson pseudo-maximum likelihood estimator. *Economics Letters 112*(2), 220–222.

Simón, H., R. Ramos, and E. Sanromá (2014). Immigrant occupational mobility: Longitudinal evidence from Spain. *European Journal of Population 30*(2), 223–255.

Solana, J. (2003). Summary of the address by mr Javier Solana, EU High Representative for Common Foreign and Security Policy to the European Parliament. *S0137/03 (Brussels), http://www. consilium. europa. eu/ueDocs/cms_Data/docs/pressdata/EN/discours/76240. pdf (accessed July 2, 2020).*

Spenkuch, J. L. (2014). Understanding the impact of immigration on crime. *American Law and Economics Review 16*(1), 177–219.

Thursby, M., R. Jensen, and J. Thursby (1991). Smuggling, camouflaging, and market structure. *The Quarterly Journal of Economics 106*(3), 789–814.

Tinbergen, J. (1962). Shaping the world economy; suggestions for an international economic policy.

UNODC (2014). The illicit drug trade through South-Eastern Europe. Technical report, United Nations Office of Drugs and Crime.

UNODC (2017). Methodology – world drug report 2017. Technical report, United Nations, New York, NY.

UNODC (2020a). World drug report, drug supply. Technical report, United Nations, New York, NY.

UNODC (2020b). World drug report, drug use and health consequences. Technical report, United Nations, New York, NY.

Yruela, M. P. and S. Rinken (2005). *La Integración de los Inmigrantes en la Sociedad Andaluza*, Volume 22. Editorial CSIC-CSIC Press.

Chapter 2

How Much Are Car Purchases Driven by Home Equity Withdrawal?

Brett A. McCully, Karen M. Pence, and Daniel J. Vine[1]

June 2018

Abstract

Previous research indicates that changes in housing wealth affect consumer spending on cars. We find that home equity extraction plays only a small role in this relationship. Consumers rarely use funds from equity extraction to purchase a car directly, even during the mid-2000s housing boom; this finding holds across three nationally representative household surveys. We find in credit bureau data that equity extraction does lead to a statistically significant increase in auto loan originations, consistent with equity extraction easing borrowing constraints in the auto loan market. This channel, though, accounts for only a tiny share of overall car purchases.

JEL Codes: D12, D14, E21, G21.

Keywords: Auto loans, auto sales, cash-out refinancing, home equity, home equity lines of credit, mortgage refinancing, motor vehicles.

2.1 Introduction

House prices in the U.S. rose dramatically from 1998 to 2006 and then plunged thereafter, bottoming out in 2011. Several studies, which we review below, have connected the changes in housing wealth during this period to the patterns in consumer spending on

[1]* Author contact information: McCully, UCLA Department of Economics; Pence and Vine, Board of Governors of the Federal Reserve System. We are grateful to Kyle Coombs and Jimmy Kelliher for extraordinary research assistance and to Ben Keys and many Federal Reserve colleagues for helpful comments. The views in this paper are the authors' alone and do not necessarily represent the views of the Board of Governors of the Federal Reserve System or its staff, or the National Institute of Child Health and Human Development of the National Institutes of Health. Brett McCully was supported by a fellowship from the Eunice Kennedy Shriver National Institute of Child Health & Human Development (T32HD007545) for part of the duration of this project.

other goods, in particular automobiles. Less is known, however, about how households deploy their home equity gains in order to purchase autos.

Narratives in the popular press suggest that it is quite common for households to use the proceeds from home equity extraction to fund auto purchases and that this practice was especially popular during the housing boom in the mid-2000s (e.g., Dash 2008, Harney 2015, Singletary 2007). The economics behind these narratives can seem a bit puzzling, however, as it is usually more cost effective for households to finance car purchases with auto loans than with home equity loans, even during housing booms. To better understand these narratives and assess how important home equity extraction actually is to funding auto purchases, in this paper we assess the two ways in which homeowners might use home equity to purchase cars. First, homeowners might use equity extraction proceeds directly to purchase cars outright. Second, households might use equity extraction proceeds indirectly to facilitate purchasing a car with an auto loan. In particular, credit constrained households might use home equity proceeds to alleviate down payment constraints in the auto loan market or to pay down high interest debt and thereby free up space in their budgets to take out an auto loan.

We find evidence that both pathways play some role in the relationship between house prices and car purchases, but neither pathway appears to have been a quantitatively important part of car purchases during the mid-2000s housing boom. We first show that very few households report purchasing cars primarily with funds from home equity lines of credit or the proceeds of cash-out refinancing, even during the housing boom years. This result is consistent across three nationally representative household surveys. We then use credit bureau data to explore whether home equity extraction indirectly supports car purchases by facilitating auto loans, and we find relatively strong evidence that this is the case. We explore the data a bit further and find that this relationship more likely reflects the role of equity extraction in easing down payment requirements in the auto loan market than an interaction between equity extraction and uncollateralized debt. But our estimates imply that the quantitative impact of home equity extraction on car purchases through the indirect auto loan channel is also quite small.

We use an event study setup in the analysis of the credit bureau data and identify the effects of home equity extraction on auto loan originations by looking for a discontinuous increase in auto loan originations shortly after equity extraction. The setup allows us to distinguish the role of equity extraction in easing auto loan credit constraints from other factors that might cause equity extraction and auto lending to move together, such as house prices and interest rates, and to assert that the relationship that we find between equity extraction and auto lending is likely causal.

Our results provide mixed support for studies in the existing literature that find that housing wealth primarily supported consumption during the 2000s by increasing the ability of households to borrow. Some patterns in the credit bureau data are consistent with this narrative, such as the stronger relationship we find between home equity extraction and subsequent auto loan origination for borrowers with low- to moderate credit scores than for other borrowers. However, individuals in the household surveys who report using home equity as the primary source of funds for purchasing a car do not appear to be particularly borrowing constrained. Because so few car purchases are funded through home equity, though, we hesitate to generalize too broadly about the implications of our findings for housing wealth and consumption.

Our results cast some doubt on the narrative that home equity extraction was an important source of funds for auto purchases during the housing boom in the mid-2000's, but they do not imply that housing wealth was inconsequential for these purchases. The wealth effects of the changes in house prices could have been large, and some of the indirect effects of home equity extraction on auto purchases that we cannot explore in the data could also have been important. In the conclusion we discuss whether households may purchase other goods and services with home equity and free up space in a household balance sheet to buy a car.

2.2 Related Literature

Our paper contributes to two literatures: (1) Studies of the relationship between house prices and consumption, and (2) studies of credit constraints in the auto loan market. Turning first to house prices and consumption, one key question in this vast literature is whether increases in house prices spur consumption primarily because households are wealthier (the wealth channel) or because lenders are willing to extend more credit to households after their house values rise (the borrowing constraints channel).[2] The studies that have examined this relationship using data from the 2000s generally conclude that borrowing constraints are the more important of the two channels (e.g., Aladangady 2017, Bhuttta and Keys 2016, Cooper 2013, and Cloyne, Huber, Ilzetzki, and Klevin 2017).

Consistent with this general finding, several studies also indicate that borrowing constraints in the mortgage market are an important part of the link between house prices and auto sales. Mian, Rao, and Sufi (2013) and Mian and Sufi (2014) find that the relationship between the changes in house prices and auto sales is strongest in zip codes where the share of the residents with high debt burdens or low incomes is high. Brown, Stein, and Zafar (2015) show that increases in house prices in the 2002 to 2006 period were associated with increases in borrowing on home equity lines of credit and auto loans; the response of auto debt to the changes in house prices was strongest for subprime borrowers. Gabriel, Iacoviello, and Lutz (2017) show that auto sales increased more between 2008

[2] Berger, Guerrieri, Lorenzoni, and Vavra (forthcoming) provide a recent treatment of the channels between house prices and consumption.

and 2010 in counties where California's foreclosure prevention programs were especially successful in stabilizing house prices after the 2007-09 recession; they attribute this result to the rise in housing wealth, which eased credit constraints.

Other studies find that auto loan originations increase when changes in mortgage finance conditions allow more households to tap their home equity. Beraja, Fuster, Hurst, and Vavra (2017) find that the drop in mortgage rates that ensued after the start of the Federal Reserve's large-scale asset purchase program resulted in the largest increase in auto purchases in MSAs with highest median home equity. They also find that auto loan originations increased more for individuals that had a cash-out refinancing than a non-cash-out refinancing. Laufer and Paciorek (2016) find that looser credit standards on mortgage refinancing are associated with an increase in auto loan originations among subprime mortgage borrowers.

Our contribution to this literature is to ask whether households who experience large house price gains subsequently use home equity extraction to fund car purchases. Other than the Beraja, Fuster, Hurst, and Vavra (2017) study, this particular question has not been investigated very thoroughly in the extant literature. We also consider whether the households who appear to purchase cars with home equity have characteristics that suggest borrowing constraints were a key factor in their choice of payment method.

Turning to the literature on credit constraints, several studies have documented that borrowing constraints are an important feature of the auto loan market, including Attanasio, Goldberg, and Kyriazidou (2008) and Adams, Einav, and Levin (2009); the latter study shows that minimum down payments matter a great deal to borrowers in the subprime auto loan market. Consistent with this result, Cooper (2010) finds in some waves of the Panel Study of Income Dynamics a positive relationship between home equity extraction and automobile costs, which include down payments on loans and leases.

A piece of empirical evidence that is commonly used to support the importance of borrowing constraints is the high contemporaneous sensitivity of auto purchases to predictable changes in income. Some studies demonstrate this sensitivity by using changes in mortgage market conditions, which affect the income that is available for non-housing purchases. For example, Agarwal et al. (2017) and DiMaggio et al. (2017) find an increase in auto loan originations after a drop in household mortgage payments due to the Home Affordable Modification Program and mortgage rate resets, respectively. DiMaggio et al. (2017) find a stronger response for homeowners with lower incomes and higher loan-to-value ratios. Other examples of predicable changes in income that appear to affect car sales contemporaneously include tax refunds (Adams, Einav, and Levin 2009, Souleles 1999); economic stimulus payments (Parker, Souleles, Johnson, and McClelland 2013); an increase in the minimum wage (Aaronson, Agarwal, and French 2012); an increase in Social Security benefits (Wilcox 1989); and expansions of health insurance (Leininger, Levy, and Schanzenbach 2010). We add to this literature by documenting that car purchases are responsive to increases in available liquidity in the form of equity extraction. We believe that we are also the first authors to explicitly link an easing of borrowing constraints in the mortgage market to an easing of borrowing constraints in the auto loan market.

2.3 Home Equity Extraction as a Source of Funds for Car Purchases

We begin by measuring the share of auto purchases that are funded directly by home equity. Using household surveys, we define a car purchase as funded directly with home equity if a respondent indicates that she bought a new or used car and that home equity was a source of funding. Our analysis is based on three surveys: The Reuters/University of Michigan Survey of Consumers (Michigan Survey), the Federal Reserve's Survey of Consumer Finances (SCF), and the Bureau of Labor Statistics' Consumer Expenditure Survey (CE). As described in Appendix A, the three surveys ask about home equity extraction and auto purchases in different ways but nonetheless show a similar relationship between these two events.

As shown in Table 1, households rarely report using home equity to purchase cars. Results from the three surveys suggest that home equity extraction funds about 1 to 2 percent of both new and used car purchases. When we run these tabulations on the SCF and CE using only data for homeowners, as renters cannot have home equity, the shares of car purchases funded with home equity are only about $\frac{1}{2}$ percentage point higher.[3] The surveys show that households typically fund new car purchases with auto loans, which finance around 70 percent of new car purchases and a somewhat smaller share of used car purchases—around 40 to 50 percent. Cash or some other source of funds are used to finance the remaining 25 percent or so of new car purchases and 50 to 60 percent of used car purchases.[4]

Although home equity appears to directly fund only a very small share of car purchases, its use might have picked up during the housing boom and then dropped off during the financial crisis. To assess this possibility, we calculated from the CE the share of car purchases funded by a home equity loan for each year between 1997 and 2012 (Figure 1). The share of cars purchased with home equity was low over the entire period; it averaged

[3]In the CE: Home equity was used by 1.0 percent of homeowners who bought a new car and 1.3 percent who bought a used car. In the SCF: Home equity was used by 2.7 percent of homeowners who bought a new car and 2.5 percent who bought a used car.

[4]The shares presented in Table 1, which are based on transaction counts, change only slightly if they are instead based on dollars spent. SCF tabulations indicate that the average purchase price was around \$25,000 for cars funded with auto loans or home equity, and \$29,000 for cars purchased with cash; the median values were even closer at \$24,000 or \$25,000 for all three funding methods.

0.7 percent both during the housing boom (1997 to 2006) and after it (2007 to 2012).[5]

There are a few reasons why it may not be surprising that the share of car buyers that report home equity as the funding source, even during the housing boom, is so low. First, personal finance professionals would generally advise against using a home equity loan to purchase a car, as these loans extend maturities beyond the lengths typically recommended for cars and thus may increase the total interest paid by consumers (Singletary, 2008; The Wall Street Journal). Second, the transaction costs of extracting home equity with a second lien or mortgage refinancing generally exceed those of originating an auto loan; doing so only makes sense if the homeowner plans to extract a lot of equity at once and use much of it for another purpose. Third, the primary advantage to using home equity rather than an auto loan to finance a car purchase—the tax deductibility of the interest for loans up to $100,000—is most likely not relevant for the approximately one-third of homeowners who end up taking the standard deduction (Poterba and Sinai 2008).[6] Finally, auto loans were an attractive financing choice during much of the housing boom period: Auto credit appears to have been widely available, and interest rates on new car loans were generally low and often heavily discounted by the manufacturers, especially for households with low credit risk.

So who uses home equity to buy cars? To answer this question and explore whether borrowing constraints are a factor, we compare the income, wealth, and credit history characteristics of households who purchase cars with home equity with those who purchase cars with auto loans or with cash or other means. We use data from the SCF for this exercise, and we limit the sample to homeowners who purchase new cars to eliminate the differences between homeowners and renters, and between new car purchasers and used car purchasers.

The comparisons, which are shown in Table 2, suggest that homeowners who report using home equity to buy a car do not appear to be lacking in terms of income, wealth, or access to credit. Among new car buyers, the table shows a clear ordering by method of funding an auto purchase: Households who use cash have the most wealth and access to credit, followed by households who use home equity and then households who use auto loans. Most of the differences among the three groups are statistically significant even with the very small sample of households who use home equity.

[5]The pattern does not appear to be substantively different for households identified in the CE as living in California, Arizona, Nevada, and Florida (states with particularly high rates of home price appreciation during the housing boom). The share of cars purchased with home equity in these states averaged 0.4 percent from 1997 to 2006 and 0.8 percent from 2007 to 2012. These tabulations are based on smaller samples than the overall shares.

[6]The Tax Cuts and Jobs Act of 2017 suspends the tax deductibility of this interest from 2018 to 2026. Under the provisions of the law, the interest on home equity loans is only tax deductible if the loan is collateralized by a loan "used to buy, build or substantially improve the taxpayer's home that secures the loan." See https://www.irs.gov/newsroom/interest-on-home-equity-loans-often-still-deductible-under-new-law for a summary of the changes.

Beginning with wealth, the median of liquid assets is $42,000 for homeowners who purchase new cars with cash, $22,000 for those who use home equity, and $10,500 for those who use an auto loan.[7] Likewise, median net worth is a bit greater than $1,000,000 for cash purchasers, nearly $600,000 for home equity purchasers, and nearly $300,000 for auto loan purchasers. The ordering of median income among the groups is the same as for wealth, but the differences are not statistically significant.

Turning to access to credit, the share of homeowners who purchase new cars and answered "yes" to the survey question "Was there any time in the past five years that you thought of applying for credit at a particular place, but changed your mind because you thought you might be turned down?" is low—only 2 percent for cash and home equity purchasers and 10 percent for auto loan purchasers. The share who answered "yes" to the survey question "In the past five years, has a particular lender or creditor turned down any request you made for credit?" is somewhat higher at 7 percent for cash purchasers, 15 percent for home equity purchasers, and 20 percent for auto loan purchasers. But the differences among the groups are not statistically significant for this measure. By both measures, homeowners who purchase new cars with home equity do not appear to be credit constrained.

Demographic characteristics that are correlated with credit access—age, education, and stockownership—similarly suggest that households who purchase new cars with home equity do not stand out as being credit constrained. Home equity purchasers are around 50 years old, on average, somewhat younger than cash purchasers (60 years old) and about the same age as auto loan purchasers. The share of home equity purchasers with a college education is about 43 percent, below the share of cash purchasers (54 percent) and about the same share as auto loan purchasers. The share of home equity purchasers that own stock is 39 percent, below the share of cash purchasers (48 percent) and above the share of auto loan purchasers (24 percent).

2.4 Home Equity Extraction as a Facilitator of Auto Loans

Although few households report directly using home equity to purchase a car, a larger number of households might indirectly use home equity to purchase a car by using the proceeds of a recent equity extraction to overcome down payment requirements or other credit constraints in the auto loan market. In this section, we use an event study set up to examine this indirect channel and estimate whether homeowners are more likely to take out an auto loan right after extracting home equity.[8]

[7]Liquid assets are defined as checking, savings, and money market accounts, and call accounts at brokerages.

[8]Other papers that have used similar event study approaches include Benmelech, Guren, and Melzer (2017) and Beraja et al. (2017).

As described in the literature review, a common way to detect the presence of borrowing constraints is to test whether auto purchases rise after a household receives a predictable boost in income. Using a similar logic, we use an event study setup to estimate the effect of home equity extraction on auto loan originations via the route of alleviating borrowing constraints in the auto loan market. Specifically, we measure the additional increase in the probability of originating an auto loan *after* home equity is extracted relative to the probability observed *before* equity is extracted.

The identification strategy of the event study setup assumes that the effects of common shocks to home equity extraction and auto loan originations—such as a wealth effect associated with a rise in house prices or a price effect associated with a change in interest rates—are equally relevant for auto loan originations before and after home equity is extracted. In contrast, when equity extraction facilitates an auto loan origination because it eases a constraint in the auto lending market, the auto loan origination must follow the extraction.

Our analysis uses credit bureau data from the Federal Reserve Bank of New York Consumer Credit Panel (CCP).[9] The panel is a randomly selected anonymized 5 percent sample of credit records from the credit bureau Equifax. The data include individuals' credit scores, debt balances, payment histories, age, and geographic location (down to the Census block level). Individuals are followed over time with quarterly snapshots of their data, although the sample is periodically refreshed so that it remains representative of all individuals with a credit record and a social security number. We use data from 1999 to 2015, and for computational ease we select a 20 percent subsample; all told, our dataset is a 1 percent sample of the universe of credit records. An observation i in our sample is the data for a given individual in a given quarter.

We construct a sample of individuals who could plausibly have extracted equity at any time in an event window that spans three quarters before and three quarters after the quarter in which we observe the individual. Those individuals are borrowers who have mortgage debt and are current on that debt throughout the event window. For each event window we drop from the sample households who appear to have purchased a new residence (as determined by a change in the census block of residence from quarter-to-quarter) or who appeared to have been property investors (as determined by the presence of more than one first lien mortgage or home equity line of credit on the credit bureau file).[10] The resulting dataset has approximately 31.5 million person-quarter observations.

Auto loan originations and home equity extractions are not directly reported in the CCP data, and so we infer these extensions of new credit from the number of open

[9]See Lee and van der Klaauw (2010) for more information on the FRBNY Consumer Credit Panel.

[10]Census blocks are the smallest unit of geography that the Census Bureau uses to tabulate decennial data. Generally blocks are quite small; in urban areas, for example, a census block often corresponds to a city block.

accounts for each borrower and their loan balances. For auto loans, we infer that a new loan was originated when the number of open auto loan accounts for a borrower increases from one quarter to the next or when the borrower's total indebtedness on non-delinquent auto loans rises.[11] As with mortgages, we do not count a balance increase on delinquent accounts as a loan origination because it may reflect overdue interest or fees being rolled into the loan balance.

To infer that a home equity extraction took place, we search our dataset for borrowers with mortgage debt in two consecutive quarters and with an increase in total mortgage debt of at least 5 percent (and at least $1,000) from the first to the second quarter. Because our dataset includes no borrowers who purchase new residences, appear to be property investors, or have a delinquent mortgage, none of the increases in mortgage balances in our dataset are associated with these activities. In addition, we flag apparent changes in the loan servicer, which can result in the reported balance on the loan dropping to zero for a quarter until the new servicer begins reporting to the credit bureau. In these cases, we replace the zero balance with the average of the balances from the prior and subsequent quarters and therefore do not record these servicing transfers as equity extractions.[12] Data limitations preclude us from following a similar procedure for servicing transfers associated with auto loans.[13]

The reason we drop borrowers who purchase residences, are property investors, or are delinquent on their mortgages from our dataset—despite the fact that borrowers in these situations may extract home equity—is that retaining these borrowers would bias our estimates downward. The source of the bias is the uncertainty present in these situations about whether increases in mortgage balances imply that home equity was extracted. Therefore, keeping borrowers with these situations in our sample and assuming that all increases in mortgage balances are not equity extractions would bias downward the relationship we estimate between equity extraction and auto loan origination. We judge the simplest solution to be to drop these households entirely.

As a baseline, we use our final dataset to estimate equation (1) and determine the likelihood that an individual takes out an auto loan, conditional on whether she has

[11] The auto loan field in credit bureau data includes auto leases as well as loans collateralized by both new and used vehicles. Our auto loan origination definition follows the definitions used in The Quarterly Report on Household Debt and Credit, available at https://www.newyorkfed.org/microeconomics/hhdc.html, and the code was generously provided by FRBNY staff. We build on their work by excluding balance increases associated with delinquent loans. The starting point for our home equity extraction code is Bhutta and Keys (2016), which is available on the American Economic Review web site at https://www.aeaweb.org/articles?id=10.1257/aer.20140040. We build upon these authors' work by excluding balance increases associated with delinquent loans and by addressing servicing transfers. We thank all these authors for their generosity in providing their code.

[12] Adjusting for servicing transfers reduces the number of equity extractions by approximately 11 percent in 1999–2001, and by 1 to 6 percent in 2002–2015.

[13] At the time this paper was written, the CCP contained individual records for each loan ("tradelines") for mortgages and HELOCs but not for auto loans.

extracted home equity recently or will do so in the near future. The dependent variable $Auto_i$ equals 1 if she originated an auto loan in the quarter associated with observation i and 0 otherwise. The independent variables include an intercept and a sequence of 7 indicator variables that correspond to the three quarters before the reference quarter associated with observation i (q = -3, -2, or -1), the reference quarter itself (q = 0), and the three quarters after the reference quarter (q = 1, 2, or 3). Each indicator variable equals 1 if the individual extracted home equity in that quarter and zero otherwise.

$$Auto_i = \alpha + \sum_{q=-3}^{3} \beta_q \times 1_i\{Extracted\ equity\ in\ quarter\ q\} + \varepsilon_i \tag{2.1}$$

We estimate equation (1) as a linear probability model and report the coefficient estimates in the first column of Table 3. As indicated by the estimate of the intercept, about 3.6 percent of individuals who did not extract home equity at any point in the relevant seven quarter window originated an auto loan in the reference quarter. The estimates of the βq coefficients, when $q < 0$, measure the additional probability that an individual takes out an auto loan in the reference quarter if they extracted home equity q quarters ago; when $q > 0$, these estimates measure the additional probability that an individual takes out an auto loan in the reference quarter if they will extract home equity q quarters in the future. Individuals are about 1.1 percentage points more likely to take out an auto loan if they extracted home equity three or two quarters ago (β_{-3} and β_{-2}). Individuals are 1.3 and 1.7 percentage points more likely to originate an auto loan if they extracted equity one quarter earlier or in the same quarter (β_{-1} and β_0). Individuals are about 1.1 percentage points more likely to originate an auto loan if they will extract home equity either 1, 2, or 3 quarters in the future (β_1, β_2, and β_3); estimates of these coefficients are essentially identical to those for β_{-3} and β_{-2}.

Estimates of all seven β coefficients are positive and statistically different from zero, indicating that individuals who have extracted home equity recently or will do so in the near future are more likely to take out an auto loan than are other individuals. This relationship may reflect factors such as rising housing wealth or low interest rates overall, which boost the likelihood of both equity extraction and auto loan origination, or characteristics of the borrowers that affect the likelihood of both activities.[14]

In addition, the β coefficients that correspond to subsets of the sample that extracted home equity during the reference quarter or one quarter before it are larger than the other β coefficients, consistent with equity extraction easing credit constraints in the auto loan market. As described earlier, our identification stems from the timing of events: if borrowers use the proceeds of home equity extraction to overcome credit constraints in the auto loan market, they cannot originate the auto loan before receiving home equity

[14]Parker (2017), for example, links the propensity of households to increase spending in response to the arrival of predictable, lump sum payments to persistent household traits.

proceeds. Assuming that other factors that affect auto loan originations do not change systematically around the point of extraction, we interpret the incremental rise in the probability of originating an auto loan after the equity extraction as the causal effect of extraction on auto loan origination. In equation (1), the β-1 coefficient estimate is 0.2 percentage point higher than that for the other βs, and the estimate of β0 is 0.6 percentage point higher, yielding a total effect of 0.8 percentage point. The difference between the β0 and β-1 estimates is statistically significant, and the implied total effect is large relative to the 4 percent unconditional probability in this sample of originating an auto loan in a typical quarter.[15] The magnitude of the effect that we measure is also similar to the increase in the probability of taking out an auto loan after equity extraction measured in Beraja, Fuster, Hurst, and Vavra (2017).[16]

One possible concern with using this event study setup with these data is that quarterly observations may be too coarse to assert that home equity extractions predated the auto loan originations when both occurred in the same quarter. The results in Beraja, Furster, Hurst, and Vavra (2017) assuage this concern. In their credit bureau data—which unlike ours are measured at a monthly frequency—auto loan originations begin to rise in the month after equity extraction, with the peak occurring two months after the extraction.

Next, we add person fixed effects to the probability model to control for each individual's innate probability of taking out an auto loan. Other than the intercept that now varies across individuals, this model, shown in equation (2), is the same as equation (1).

$$Auto_i = \alpha_i + \sum_{q=-3}^{3} \beta_q \times 1_i\{Extracted\ equity\ in\ quarter\ q\} + \varepsilon_i \tag{2.2}$$

The coefficient estimates from this specification are shown in the second column of Table 3. The estimates of β_{-3}, β_{-2}, β_1, β_2, and β_3 are around 0.4 percentage point, β_{-1} is 0.6 percentage point, and β_0 is 1.0 percentage point. These β coefficient estimates are all about 0.7 percentage point below the corresponding estimates in equation (1), a comparison that indicates that individual heterogeneity explains much of the correlation between auto loan origination and home equity extraction. However, even with the addition of the fixed effects, individuals are still 0.8 percentage point more likely to originate an auto loan in the reference quarter if they extract home equity in that quarter or the one before it, and this increase remains statistically significant. So the conclusion that equity extraction eases borrowing constraints in the auto loan market is robust to the inclusion of person fixed effects.

In equation (3) we add year fixed effects to the model in equation (2) to capture omitted factors that vary over time but not individuals; examples of these factors might

[15]Specifically, an F-test rejects at the 1 percent level the hypotheses that β_0 and β_1 are equal.
[16]See Figure 7 in Beraja, Fuster, Hurst, and Vavra (2017).

include the level of interest rates or the national unemployment rate, which affect both equity extraction and auto loan originations. Coefficient estimates from this specification are in the third column of Table 3.

$$Auto_i = \alpha_i + \sum_{q=-3}^{3} \beta_q \times 1_i\{Extracted\ equity\ in\ quarter\ q\} + \sum_{y=2001}^{2014} \pi_y \times 1_i\{Year = y\} + \varepsilon_i \tag{2.3}$$

The estimates of the β_{-3}, β_{-2}, β_1, β_2, and β_3 coefficients are even lower in this specification than in equation (2)—by around 0.2 percentage point. Although the coefficient estimates are still statistically significantly different from zero, for practical purposes home equity extraction affects the probability of originating an auto loan in this specification only if the extraction occurs in the reference quarter or the quarter before it. The probability of originating an auto loan in the reference quarter is 0.7 percentage point higher if equity is extracted in the same quarter than if it is extracted in the next quarter, and the probability is 0.3 percentage point higher if equity is extracted one quarter earlier.[17]

Finally, in equation (4) we add the one year change in a house price index for the borrower's Zip code, ΔHPI_i, to the model in equation (3). Changes in local house prices can vary considerably across the country and therefore are only partly captured by the year fixed effects. The Zip code house price indexes are from CoreLogic, and we are able to match these indexes to borrowers for 72 percent of the borrowers in the sample.[18] We include this specification to take into account the tendency of some households to make a number of home price appreciation-related financial decisions at one time. If paying attention is costly, for example, households might react to an increase in house prices by extracting equity and originating an auto loan in the same quarter. In this case, it is the fixed cost of paying attention rather than the presence of borrowing constraints that explains the pattern of the relevant beta coefficient estimates in equations (1) through (3). Coefficient estimates from this house price augmented specification are in the fourth column of Table 3.

[17] We also ran a specification with quarter-year fixed effects; the results were essentially the same as the year fixed effects specification.

[18] This match rate is consistent with that found in Bhutta and Keys (2016). See p. 1749 of that paper for more discussion.

$$Auto_i = \alpha_i + \sum_{q=-3}^{3} \beta_q \times 1_i\{Extracted\ equity\ in\ quarter\ q\} + \gamma \times \Delta HPI_i + \sum_{y=2001}^{2014} \pi_y \times 1_i\{Year = y\} + \varepsilon_i \quad (2.4)$$

The estimate of γ indicates that the association of regional house price changes with auto loan originations is statistically significant but very small; the 0.00007 coefficient estimate means that even a fairly large one year house price increase of 19 percent (the 95th percentile of one year house price increases in our sample) is associated with an increase in the probability of originating an auto loan of only 0.1 percentage point. In a more flexible nonlinear specification (not shown), we allow γ to vary across six increments of house price increases and similarly find that living in a Zip code with the largest increase—a 10 percent or greater increase from the previous year—is associated with only a 0.1 percentage point increase in the probability of originating an auto loan.

Importantly, the estimates of the sequence of β coefficients are essentially unchanged in this specification relative to equation (3). As before, the coefficient estimates imply that borrowers are 0.7 percentage point more likely to originate an auto loan in the reference quarter if they extract home equity in the same quarter and 0.3 percentage point more likely to do so if they extract equity one quarter earlier. Similarly, characterizing house prices with the nonlinear transformation described above has little effect on the β coefficient estimates, and the same is true if we use three year changes in the regional house price indexes in place of one year changes.

The various robustness checks support our conclusion that the rise in auto loan originations that occurs during and shortly after a home equity extraction stems from an easing of credit constraints.

2.5 Additional Evidence of Borrowing Constraints in the Auto Loan Market

If home equity extraction boosts the likelihood of taking out an auto loan because it eases credit constraints in the auto loan market, we would expect the relationship to be stronger for borrowers with lower credit scores. We look for this corroborating evidence by adding variables to the event study probability model that allow the intercept and coefficients on the equity extraction time indicators to vary for six credit score categories, indexed by c.[19] As in equation (3), this specification includes person and year fixed effects.

$$Auto_i = \alpha_i + \sum_{c=1}^{6} \alpha_c +$$
$$\sum_{c=1}^{6} \sum_{q=-3}^{3} \beta_{q,c} \times 1_i\{Extracted\ equity\ in\ quarter\ q\} \times 1_i\{Credit\ category = c\}$$
$$+ \sum_{y=2001}^{2014} \pi_y \times 1_i\{Year = y\} + \varepsilon_i \quad (2.5)$$

The β_{qc} coefficient estimates are graphed in Figure 2.[20] The probability of taking out an auto loan in the reference quarter is higher for borrowers in all of the credit score groups who extract home equity in the same quarter or one quarter earlier, but the magnitudes of the increases vary substantially. These probabilities are shown in the second column of Table 4. For comparison, the first column shows the probability that a borrower from each group originates an auto loan in the reference quarter if they extract home equity during the next quarter. As in the earlier exercises, this probability represents the rate at which borrowers take out an auto loan if they extract equity but do not face borrowing constraints. To gauge the contribution of the role of home equity in easing credit constraints for each group, column 3 shows the percent increase in the probability of originating an auto loan associated with having extracted equity before or during the reference quarter as opposed to after it.

Individuals with subprime credit scores who extract home equity after the reference quarter have a 5.0 percent probability of originating an auto loan in the reference quarter. If these individuals instead extract equity during the reference quarter or the quarter before it, that probability is 7.1 percent, which represents a 42 percent increase in the probability of originating an auto loan. In contrast, individuals with the highest (ultra-prime) credit scores who extract home equity after the reference quarter have a 5.8 percent probability of originating an auto loan in the reference quarter, and that likelihood only edges up to 5.9 percent if the home equity extraction instead occurs during the same quarter or one quarter. In relative terms, borrowing constraints have barely any impact on the rate at which this group originates auto loans. For individuals with middle credit scores, extracting equity before or during the reference quarter increases the probability of

[19]The credit scores on the CCP are Equifax 3.0 risk scores, which range from 280 to 850 and are roughly comparable to FICO credit scores. Our score ranges are as follows: Deep subprime, 280–579; subprime, 580–619; near prime, 620 to 659; prime, 660 to 699; super-prime, 700 to 759; and ultra-prime, 760 to 850. We thank Hank Korytkowski for help in determining these ranges. Consumer Financial Protection Bureau (2012) provides a comparison of the different types of credit scores.

[20]We omit borrowers with "deep subprime" credit scores from Figure 2 because the very few households in this category who own homes likely have credit scores too low to qualify for additional mortgage credit.

originating auto loans between 22 and 27 percent. The larger effect observed for subprime individuals relative to other groups corroborates our conclusion that credit constraints underlie the relationship between equity extraction and auto loan originations identified by the event study.

Although our data do not reveal much about how home equity extraction eases credit constraints in the auto loan market, one theory we can test is whether borrowers who extract home equity appear to use the proceeds to pay down high interest consumer debt. Bhutta and Keys (2016) show that credit card debt decreases only slightly after a home equity extraction, on average, but the decreases are larger and more persistent for individuals with lower credit scores. Such a maneuver could make a household a better credit prospect by reducing its credit utilization rate (which counts toward a borrower's credit score) and its debt service relative to income (which might be a factor in auto loan underwriting).[21] Independent of the lender's determination, the borrower might feel a greater capacity to take out an auto loan after paying down higher interest debt.

To look for evidence of consumer debt paydown, we construct an indicator variable CC_Pay_i that equals 1 if the individual pays down half or more of the existing uncollateralized consumer debt in the quarter associated with observation i, and it is set to 0 otherwise.[22] We then assess whether individuals who took this action are more likely than other equity extractors to purchase cars. The exercise is shown as equation (6), which includes a term that interacts CC_Pay_i with an indicator of whether equity was extracted in the reference quarter.[23] A positive and significant estimate of η would suggest that consumer debt paydown is part of the relationship between home equity extraction and auto loan originations.

$$Auto_i = \alpha_i + \sum_{c=1}^{6} \alpha_c + \sum_{c=1}^{6} \sum_{q=-3}^{3} \beta_{q,c} \times 1_i\{Extracted\ equity\ in\ quarter\ q\} + \delta \times CC\_Pay_i$$

$$+ \eta \times CC\_Pay_i \times 1_i\{Extracted\ equity\ in\ quarter\ q=0\} + \sum_{y=2001}^{2014} \pi_y \times 1_i\{Year = y\} + \varepsilon_i$$

The coefficient estimates for equation (6) (column 5, Table 3) show no detectable role for consumer debt paydown in the relationship between equity extraction and auto

[21] Replacing \$10,000 in credit card debt with \$10,000 in mortgage debt, for example, would net a household \$250 in savings per month, assuming a thirty year fixed rate mortgage at a rate of 4 percent and a monthly minimum payment requirement of 3 percent on the credit card debt.

[22] Uncollaterized consumer debt is defined as total balances on credit cards issued by banks and consumer finance companies and on retail cards as well as uncollateralized consumer installment loans.

[23] Specifications that interact credit card paydown and home equity extraction with credit score category also find essentially no relationship between these factors. We also established that our results are robust to using an indicator variable for whether the individual paid down at least a fourth of their existing uncollateralized debt in the quarter associated with observation i.

lending. The estimate of δ indicates the likelihood that an individual takes out an auto loan in the reference quarter rises 0.5 percentage point if she pays down uncollateralized debt in the same quarter, but the miniscule and insignificant estimate of η suggests this probability is nearly unaffected by whether she also extracts home equity in that quarter. More flexible specifications that interact leads and lags of credit card paydown with leads and lags of home equity extraction also find essentially no relationship between these factors.

This result is not surprising, as the existing literature suggests that the constraint most likely eased by equity extraction is down payment requirements. These studies identify down payments as a major credit constraint in the auto lending market (Adams, Einav, and Levin 2009) and find a relationship between equity extraction and increased spending on auto loan down payments (Cooper 2010).

2.6 How Important is Home Equity Extraction in the Auto Loan Market?

Having established that home equity extraction has a statistically significant effect on auto loan originations, we now ask whether equity extraction plays a quantitatively important role in aggregate auto loan originations. To begin, we first estimate in our data that around 3 million households extracted home equity each year in 2001 and 2002, and 4 to 5 million households extracted equity annually in the peak housing boom years from 2003 to 2006.[24] These estimates are consistent with the estimates of Canner, Dynan, and Passmore (2002) for the 2001 to 20002 period.[25]

To calculate the effect of these equity extractions on car purchases, we apply to the extraction volumes the coefficients from our preferred specification in equation (3), which imply that home equity extraction raises the probability of an auto loan origination by 0.9 percentage point. (This is the incremental probability of an auto loan origination in a given quarter associated with households who extract equity in the same quarter ($\beta_0 - \beta_1 = 0.65$) plus the incremental probability associated with those who extract equity in the preceding quarter ($\beta_{-1} - \beta_1 = 0.25$)).[26] Applying this 0.9 percentage point effect to

[24] Our sample is a 1 percent extract of all credit bureau records, so we multiply our sample estimates by 100 to scale them up to the national level. One complication is that about three-fourths of the home equity extractions that we observe in the credit bureau records are co-signed loans, in which case the loan observed for a borrower in the sample also belongs to another borrower that is most likely outside the sample. To account for this, we assign each co-signed loan a weight of 0.5 when we scale up the estimates.

[25] Canner, Dynan, and Passmore (2002) find that 4.6 percent of a sample of 2,240 homeowners surveyed by the Michigan Survey of Consumers had engaged in a cash-out refinancing in 2001 or the first half of 2002. They note that multiplying that fraction by the total number of U.S. homeowners at that time leads to an estimated 4.9 million cash-out refinancings in 2001 and the first half of 2002 (an annual rate of around 3 million).

[26] We rounded these estimates to 0.7 and 0.3, respectively, in the earlier discussion of the regression

4.5 million equity extractions during the peak housing boom years suggests that home equity extraction facilitated about 40,000 auto loans per year during this period. By comparison, the number of auto loan originations in recent years has varied from a low of around 15 million in 2009 to a high of around 30 million in 2016.[27] This comparison suggests that home equity extraction was likely not a quantitatively important factor in total auto loan originations or in the changes in loan originations during this period.

2.7 Conclusions

In this paper, we demonstrate that home equity extraction does not appear to be the direct source of funding for many car purchases. Estimates from three nationally representative surveys indicate that very few households purchase cars directly with home equity. Further, the share of those who report doing so does not appear to vary with the housing cycle.

However, home equity extraction is associated with an increase in auto loan originations. Using an event study framework with credit bureau data, we show that home equity extraction increases the likelihood of originating an auto loan in a statistically significant and causal way. We also show that this increase is distinct from the effects of other factors that cause equity extractions and auto loan originations to move together, such as the changes in house prices and interest rates, and that the effect of home equity extraction on auto loan originations is more pronounced for borrowers with lower credit scores. Our results suggest that home equity extraction increases auto loan originations by easing down payment and other credit constraints in the auto loan market. In contrast, we find no evidence that equity extraction increases auto loan originations by allowing households to pay down high interest debt and thereby free up space in their budgets for auto loan payments. Nonetheless, when we put the effects we estimate into the context of the U.S. auto loan market, the number of additional auto loan originations in recent years that we can attribute to home equity extraction is very small.

Our results cast doubt on the narrative that home equity extraction was an important source of funds for auto purchases during the housing boom in the mid-2000's, but they do not imply that housing wealth was inconsequential for these purchases. At least two other (not mutually exclusive) channels contribute to the relationship between auto purchases and home equity. First, households are wealthier when their homes increase in value, and their demand for cars should also increase. Rising housing wealth may have boosted car

results.

[27] See Consumer Financial Protection Bureau, "Origination Activity," https://www.consumerfinance.gov/data-research/consumer-credit-trends/auto-loans/origination-activity/, and Equifax, "Quarterly U.S. Consumer Credit Trends," https://investor.equifax.com/~/media/Files/E/Equifax-IR/reports-and-presentations/events-and-presentation/consumer-credit-trends-report-q1–2017-revised-pdf.pdf.

purchases considerably, even if these households did not report purchasing cars directly with home equity; different types of wealth are, to some extent, interchangeable. For example, paying for other goods and services with home equity may free up balance sheet space to purchase a car with cash or an auto loan.

Second, home equity might indirectly facilitate auto loans if lenders are more willing to extend credit to households in neighborhoods with rising house prices. Home equity is typically not considered directly in the underwriting of auto loans, but lenders may take into account local economic conditions, which can be correlated with house prices. Alternatively, lenders may have an easier time raising capital in areas of the country that are booming. Households in these markets may also be more likely to retain their good credit standing when their income is disrupted, because they can more easily refinance their mortgages or sell their homes. Ramcharan and Crowe (2013) show that peer-to-peer lenders were less willing to extend unsecured credit to homeowners in areas with declining house prices; a similar dynamic may occur in the auto credit market, although we are not aware of any research on this topic.

Appendix A: Survey Data

A.1 The University of Michigan Surveys of Consumers (Michigan survey)

The Michigan survey data come from a special module that the Federal Reserve has sponsored three times per year since 2003. Survey respondents are asked if they purchased a car in the previous six months, and if so, whether they borrowed money to purchase the car or paid cash. If the answer is "cash," respondents are asked whether the source of the cash was savings or investments, a home equity loan, a mortgage refinancing, or "somewhere else."[28]

Respondents can cite multiple sources of the cash, although this is rare. We define the car purchase as a home equity extraction if the respondent identifies a home equity loan or mortgage refinancing as the source of the cash. We define the purchase as an auto loan if the respondent indicates that a car was purchased with borrowed money. We define all other purchases as cash/other. The data span the 2003 to 2014 period and include 2,388 purchases of new and used cars.

[28] According to the Michigan survey staff, some respondents who purchase autos with home equity appear to consider these purchases as funded with "borrowed" money rather than "cash." If so, the survey instrument will miss some car purchases funded by home equity extraction. The survey staff catch many of these instances and recode the answers as cash/home equity. We do not think that this aspect of the question structure leads to a significant understatement of home equity funded purchases because the Michigan results are in line with the results from the other two surveys, which have different question structures.

A.2 Consumer Expenditure Survey (CE)

In the CE, households are asked about the vehicles that they currently own. We focus on cars purchased in the survey year. For each car owned, households are asked whether any portion of the purchase price was financed.[29] If so, they are asked whether the source of credit was a home equity loan. Households are not asked if the car was purchased with the proceeds from a cash-out refinancing, and so we will miss these purchases.

We define the purchase as a home equity extraction if the respondent identifies a home equity loan as a source of credit. We define the purchase as an auto loan if the respondent financed the purchase but does not indicate they used a home equity loan. We define all other purchases as cash/other. The data cover the 1997 to 2012 period and include 28,290 car purchases.

A.3 Survey of Consumer Finances (SCF)

In the SCF, like in the CE, households are asked about the cars that they own at the date of the interview. We focus on cars that were likely purchased recently. For used cars, the date of purchase is known from a survey question, and we select cars purchased during the survey year. For new cars, we must deduce the date of purchase, because the survey asks only about the model year of the car. We define a new car as recently purchased if its model year corresponds to the survey year or the subsequent year. Most new car purchases covered by this definition will have occurred during the survey year, although some of these purchases will have occurred during the previous calendar year. The reason is that new models are introduced during the previous calendar year and are not fully phased out until the subsequent calendar year. For the same reason, our definition will miss the small volume of new cars still being sold from earlier model years.[30]

The definitions described above yield a sample of car purchases from the SCF that have occurred mostly within a year of the interview. Taking advantage of this relatively short lookback window, we match households' recent car purchases to the answers from separate questions asked about outstanding auto loan balances and recent activities with home mortgages. Unlike the CE, the SCF does not ask households whether their cars were purchased with home equity, and so we infer these purchases when an SCF respondent

[29] The CE asks households a separate set of questions about the vehicles they purchased during the reference period. Our analysis is based on the set of questions about vehicles owned (in the EOVB files) because these data include questions about how the purchases were financed.

[30] In the 2013 SCF, for example, our definition would include new cars from the 2013 or 2014 model years, about 75 to 80 percent of which likely occurred in 2013 and 20 to 25 percent in 2012. Our definition misses new cars from the 2012 or earlier model years that were purchased in 2013, a volume that is likely only about 3 percent of the new car sales in our sample. These estimates are based on monthly sales by model year from JD Power and Associates and are adjusted to reflect the fact that SCF interviews are conducted from April of the survey year to the following February. Dettling et al. (2015) document that auto sales in the SCF line up well with the NIPA aggregates once the timing and model year issues are taken into account.

both appears to have purchased a car recently and reports having used the proceeds from a recently originated cash-out refinancing, second or third lien, or HELOC to buy a car.[31] If a household does not appear to have used home equity but does report having an auto loan outstanding, we assume the car was purchased with an auto loan. All other purchases are defined as cash/other.

One potentially important consequence of using the definitions described above is that households who buy the newest models early in the model year are likely overrepresented in our SCF sample of new car purchases. And, as noted earlier, we also miss a few purchases of older car models. All told, these factors may bias upward some of the sample statistics on new car buyers, such as average income and wealth, because new car prices decline over the course of the model year (Aizcorbe, Bridgman, and Nalewaik, 2009) and can drop when newer models are introduced.[32] These price dynamics suggest that households who buy new cars immediately upon the model release are likely more affluent than those who purchase later in the model year.

We use data from the 2004, 2007, 2010, and 2013 surveys, which include 3,929 purchases of new and used cars.

Tables and Figures

Table 2.1: Percent of Cars Purchased with Each Source of Funds

	New cars			Used cars		
Funded with:	Michigan Survey	SCF	CE	Michigan Survey	SCF	CE
Home equity	1	2.3	0.9	2.6	1.6	0.6
Auto loan	72	69	75	53	40	44
Cash/other	27	28	24	44	58	56
Memo:						
N	830	1,864	14,385	1,062	2,118	36,718

Notes: Table excludes leases. Estimates from the Michigan Survey are based on data from 2003 to 2014. Estimates from the SCF are based on data from 2004, 2007, 2010, and 2013. Estimates from the CE are based on data from 1997 to 2012. Figures in the table are calculated with sample weights provided by each survey.

[31] We consider the origination of a cash-out refinancing or second lien to be recent if it occurred in the survey year or in the year prior. We include the prior year because, as described earlier, our sample of recent vehicle purchases likely includes some cars purchased in the previous year, and because there may be a lag between the cash-out refinance and the purchase of the car. We assume that a HELOC funded a recent car purchase if the proceeds of the most recent draw were used for a car. The SCF does not ask when that draw took place; depending on the timing, our definition could either understate or overstate the share of vehicle purchases funded with HELOCs.

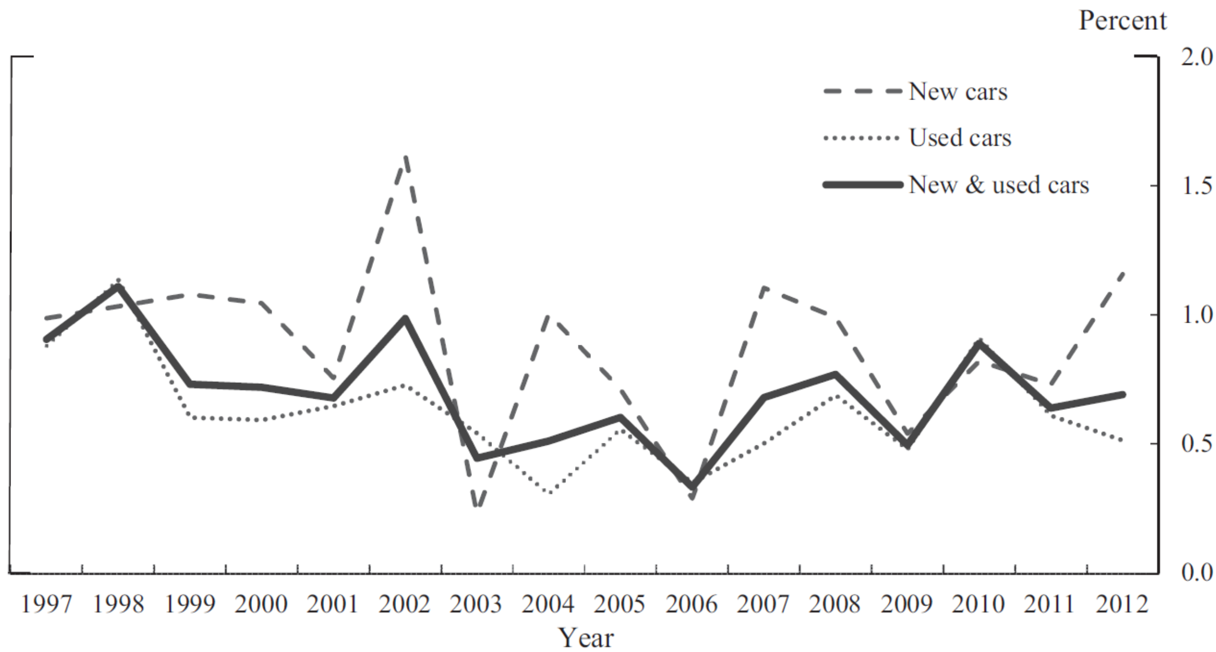
[32] he SCF and CE samples also miss vehicles purchased during the calendar year but sold (or scrapped) before the date of the survey. We assume, given our short lookback period, that this bias is small.

Table 2.2: Summary Statistics for Homeowners who Buy New Cars

Summary statistic:	Method of funding new car purchase: Cash/other	Home equity	Auto loans
Median liquid assets (2013 dollars)	42,134*	21,909	10,500**
Median net worth (2013 dollars)	1,057,283**	598,906	291,000**
Median family income (2013 dollars)	129,204	114,026	99,909
Turned down for credit in past five years (%)	7	15	20
Did not apply b/c worried turned down for credit (%)	2	2	10**
Avg. age of household head (Years)	60**	50	48
College graduate (%)	54	43	43
Own stock directly (%)	48	39	24
Memo:			
N	992	29	686

Notes: Authors' calculations from Survey of Consumer Finances data (2004, 2007, 2010, and 2013). Figures are calculated with sample weights. Figures in 2013 dollars are calculated with the Consumer Price Index from the Bureau of Labor Statistics. Asterisks denote the statistical significance of the summary statistics in each column from the estimates for those who purchase a car with home equity; *denotes the 5 percent level and ** is the 1 percent level. Statistical significance is based on standard errors bootstrapped with 999 replicates drawn in accordance with the SCF sample design and adjusted for imputation uncertainty.

Figure 2.1: Share of Cars Purchased with a Home Equity Loan



Note: Authors' calculations based on data from the Consumer Expenditure Survey. Figure shows the percent of car purchases for which the respondent cites a home equity loan as a source of financing. Shares are calculated with sample weights.

Table 2.3: Coefficient Estimates for Equations (1) through (4) and (6)
Dependent variable: Indicator for originating an auto loan in the reference quarter

	Coefficients/standard errors for specifications				
Variables	(1)	(2)	(3)	(4)	(5)
Constant	0.036**	0.037**	0.049**	0.048**	0.049**
	(0.00003)	(0.00003)	(0.00021)	(0.00020)	(0.00021)
Extracted before the reference quarter					
β_{-3} ($q = -3$)	0.011**	0.004**	0.003**	0.003**	0.003**
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
β_{-2} ($q = -2$)	0.011**	0.004**	0.002**	0.002**	0.002**
	(0.0003)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
β_{-1} ($q = -1$)	0.014**	0.006**	0.005**	0.005**	0.004**
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Extracted during the reference quarter					
β_0 ($q = 0$)	0.018**	0.010**	0.009**	0.009**	0.009**
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Extracted after the reference quarter					
β_1 ($q = 1$)	0.011**	0.004**	0.002**	0.002**	0.002**
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
β_2 ($q = 2$)	0.011**	0.004**	0.002**	0.002**	0.002**
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
β_3 ($q = 3$)	0.011**	0.004**	0.002**	0.002**	0.002**
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Annual change in house prices in borrower's zip code				0.00007**	
				(0.000006)	
Uncollateralized debt paydown during the reference quarter					0.005**
					(0.0002)
Uncollateralized debt paydown & extract during the reference quarter					−0.000003
					(0.0014)
Other control variables:					
Borrower fixed effects?	N	Y	Y	Y	Y
Year fixed effects?	N	N	Y	Y	Y
Memo: R^2	0.0005	0.07	0.07	0.07	0.07
N	31,700,439	31,700,439	31,700,439	22,982,245	31,700,439

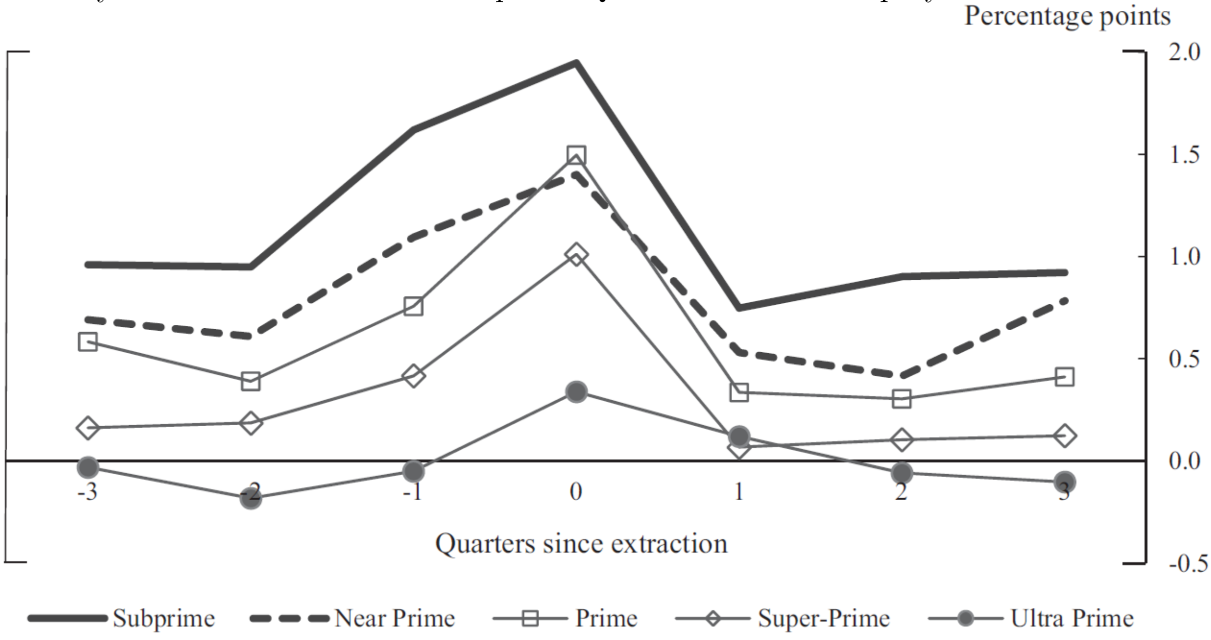
Notes: Authors' calculations from merged FRBNY CCP and CoreLogic data. Each column in the table shows the estimated coefficients from an ordinary least squares regression. Robust standard errors are displayed in parentheses underneath each coefficient. The dependent variable in each regression is an indicator variable for whether the individual originated an auto loan in the quarter of the observation. The regressions with individual fixed effects include 1,219,680 such intercepts. Two asterisks (**) denote that coefficient estimates are statistically significantly different from zero at the 1 percent level.

Table 2.4: Probability of Auto Loan Origination by Credit Score Group and Timing of Home Equity Extraction

Credit score group (risk score range)	Timing of home equity extraction: After the reference quarter (1)	Timing of home equity extraction: Before or during the reference quarter (2)	Memo: Percent change in auto loan originations from extraction before or during vs. after the reference quarter (3)
Subprime (580–619)	0.050	0.071	41%
Near prime (620–659)	0.054	0.069	27%
Prime (660–699)	0.057	0.073	28%
Super prime (700–759)	0.057	0.070	23%
Ultra prime (760–850)	0.058	0.058	1%

Notes: Authors' calculations from equation (5) estimated with FRBNY CCP data. The first column in the table shows estimates of $\alpha_c + \beta_{1c}$ for credit score groups 2 through 6. The second column shows estimates of $\alpha_c + (\beta_{-1c} - \beta_{1c}) + (\beta_{0c} - \beta_{1c})$ for credit score groups 2 through 6. The estimates in the first and second columns also include the intercept generated by STATA's areg procedure. The third column is the percent change of the second column from the first column. The credit score is the Equifax 3.0 risk score.

Figure 2.2: Effect of Home Equity Extraction on the Probability of Originating an Auto Loan by Borrower Credit Risk Group and Quarter relative to Equity Extraction



Note: Authors' calculations based on data from the Equifax Consumer Credit Panel. Figure shows the βqc coefficient estimates from equation (4).

Literature Cited

Aaronson, Daniel, Sumit Agarwal, and Eric French. (2012) "The Spending and Debt Response to Minimum Wage Hikes." American Economic Review, 102(7), 3111-3139.

Adams, William, Liran Einav, and Jonathan Levin. (2009) "Liquidity Constraints and Imperfect Information in Subprime Lending." American Economic Review, 99(1), 49-84.

Agarwal, Sumit, Gene Amromin, Zahi Ben-David, Souphala Chomsisengphet, Tomasz Piskorski, and Amit Seru. (2017) "Policy Intervention in Debt Renegotiation: Evidence from the Home Affordable Modification Program." Journal of Political Economy. 125(3), 654–712.

Aizcorbe, Ana, Benjamin Bridgman, and Jeremy Nalewaik. (2009) "Heterogeneous Car Buyers: A Stylized Fact." Economics Letters, 109(1), 50-53.

Aladangady, Aditya. (2017) "Housing Wealth and Consumption: Evidence from Geographically Linked Microdata." American Economic Review, 107(11), 3415-3446.

Attanasio, Orazio, Pinelopi Goldberg, and Ekaterini Kyriazidou. (2008) "Credit Constraints in the Market for Consumer Durables: Evidence from Micro Data on Car Loans." International Economic Review, 49(2), 401-436.

Benmelech, Efraim, Adam Guren, and Brian T. Melzer. (2017) "Making the House a Home: The Stimulative Effect of Home Purchases on Consumption and Investment." NBER Working Paper No. 23570.

Berger, David, Veronica Guerrieri, Guido Lorenzoni, and Joseph Vavra. (Forthcoming) "House Prices and Consumer Spending." Review of Economic Studies.

Beraja, Martin, Andreas Fuster, Erik Hurst, and Joseph Vavra. (2017) "Regional Heterogeneity and Monetary Policy." NBER Working Paper No. 23270.

Bhutta, Neil, and Benjamin J. Keys. (2016) "Interest Rates and Home Equity Extraction During the Housing Boom." American Economic Review, 106(7), 1742-1774.

Brown, Meta, Sarah Stein, and Basit Zafar. (2015) "The Impact of Housing Markets on Consumer Debt: Credit Report Evidence from 1999 to 2012." Journal of Money, Credit, and Banking, 47(S1), 175-213.

Canner, Glenn, Karen E. Dynan, and Wayne Passmore. (2002) "Mortgage Refinancing in 2001 and Early 2002." Federal Reserve Bulletin, 88, 469-481.

Cloyne, James, Kilian Huber, Ethan Ilzetzki, and Henrik Kleven. (2017) "The Effect of House Prices on Household Borrowing: A New Approach." NBER Working Paper No. 23861.

Consumer Financial Protection Bureau. (2012) "Analysis of Differences between Consumer- and Creditor-Purchased Credit Scores." http://files.consumerfinance.gov/f/201209_Analysis_Differences_Consumer_Credit.pdf

Cooper, Daniel. (2010) "Did Easy Credit Lead to Overspending? Home Equity Borrowing and Household Behavior in the Early 2000s." Federal Reserve Bank of Boston

Public Policy Discussion Paper No. 09-7.

Cooper, Daniel. (2013) "House Price Fluctuations: The Role of Housing Wealth as Borrowing Collateral." Review of Economics and Statistics, 95(4), 1183-1197.

Dash, Eric. (2008) "Auto Industry Feels the Pain of Tight Credit," New York Times, May 27, 2008, A-1. http://www.nytimes.com/2008/05/27/business/27auto.html

Dettling, Lisa J., Sebastian J. Devlin-Foltz, Jacob Krimmel, Sarah J. Pack, and Jeffrey P. Thompson. (2015) "Comparing Micro and Macro Sources for Household Accounts in the United States: Evidence from the Survey of Consumer Finances." Finance and Economics Discussion Series 2015-086. Washington: Board of Governors of the Federal Reserve System, http://dx.doi.org/10.17016/FEDS.2015.086.

DiMaggio, Marco, Amir Kermani, Benjamin J. Keys, Tomasz Piskorski, Rodney Ramcharan, Amit Seru, and Vincent Yao. (2017) "Interest Rate Pass-Through: Mortgage Rates, Household Consumption, and Voluntary Deleveraging." American Economic Review, 107(11), 3550-3588.

Gabriel, Stuart A., Matteo M. Iacoviello, and Chandler Lutz. (2017) "A Crisis of Missed Opportunities? Foreclosure Costs and Mortgage Modification During the Great Recession." Manuscript.

Greenspan, Alan, and James Kennedy. (2008) "Sources and Uses of Equity Extracted from Homes." Oxford Review of Economic Policy, 24(1), 120-144.

Harney, Kenneth. (2015) "Boom in Equity Allows Homeowners to Cash in and Even Cash Out." The Washington Post, September 29, 2015. http://www.washingtonpost.com/realestate/boom-in-equity-allows-homeowners-to-cash-in-and-even-cash-out/2015/09/28/7cfb940c-660d-11e5-9223-70cb36460919_story.html

Laufer, Steven and Andrew Paciorek. (2016) "The Effects of Mortgage Credit Availability: Evidence from Minimum Credit Score Lending Rules." Finance and Economics Discussion Series 2016-098. Washington: Board of Governors of the Federal Reserve System, https://doi.org/10.17016/FEDS.2016.098.

Lee, Donghoon, and Wilbert van der Klaauw. (2010) "An Introduction to the FRBNY Consumer Credit Panel." Federal Reserve Bank of New York Staff Reports No. 479.

Leininger, Lindsay, Helen Levy, and Diane Schanzenbach. (2010) "Consequences of SCHIP Expansions for Household Well-Being." Forum for Health Economics and Policy, 13(1), 1-32.

Mian, Atif, Kamalesh Rao, and Amir Sufi. (2013) "Household Balance Sheets, Consumption, and the Economic Slump." Quarterly Journal of Economics, 128(4), 1687-1726.

Mian, Atif and Amir Sufi. (2014) "House Price Gains and U.S. Household Spending from 2002 to 2006." NBER Working Paper No. 20152.

Parker, Jonathan A. (2017) "Why Don't Households Smooth Consumption? Evidence from a 25 Million Dollar Experiment." American Economic Journal: Macroeconomics, 9(4), 153–183.

Parker, Jonathan A., Nicholas S. Souleles, David S. Johnson, and Robert McClelland. (2013) "Consumer Spending and the Economic Stimulus Payments of 2008." American Economic Review, 103(6), 2530-2553.

Poterba, James, and Todd Sinai. (2008) "Tax Expenditures for Owner-Occupied Housing: Deductions for Property Taxes and Mortgage Interest and the Exclusion of Imputed Rental Income." American Economic Review, 98(2), 84-89.

Ramcharan, Rodney, and Christopher Crowe. (2013) "The Impact of House Prices on Consumer Credit: Evidence from an Internet Bank." Journal of Money, Credit, and Banking, 45(6), 1085-1115.

Singletary, Michelle. (2007) "Loan Loser: Home-Financing a Car." The Washington Post, March 18, 2007. http://www.washingtonpost.com/wp-dyn/content/article/2007/03/17/AR2007031700104.html.

Souleles, Nicholas S. (1999) "The Response of Household Consumption to Income Tax Refunds." American Economic Review, 89(4), 947-58.

The Wall Street Journal. "How to Finance an Auto Purchase" in How-to Guide: Personal Finance available at http://guides.wsj.com/personal-finance/buying-a-car/how-to-finance-an-auto-purchase/

Wilcox, David. (1989) "Social Security Benefits, Consumption Expenditures, and the Life Cycle Hypothesis." Journal of Political Economy, 97(2), 288-304.

Chapter 3

The Effect of an Early-Childhood and Fertility Intervention on Firm and Farm Productivity

Tania Barham (CU-Boulder), Randall Kuhn (UCLA), Brett A. McCully (UCLA), Patrick Turner (Notre Dame)[1]

Abstract

While health improvements are thought to be a necessary precondition for economic growth, few contemporary studies have estimated the effects of population health interventions on long-run productivity growth. We explore the long-run effects of the quasi-randomized Matlab Maternal and Child Health and Family Planning (MCH-FP) program introduced in a rural area of Bangladesh from 1977-1988 on firm and farm profitability 35 years later. Using a rich data set including two followup surveys with unusually low rates of attrition, we observe an emerging association between the program and agricultural productivity. In 1996, when children born during the program period were still young, we observed no effects. However, a large effect emerges by 2012, including a 29% increase in productivity per acre in treatment versus control area and a 43% relative increase in the adoption of high-yield seeds. We also find that the program increased the profitability and sophistication of household enterprises.

[1]We thank seminar participants at UCLA for helpful comments. McCully acknowledges financial support from CCPR's Population Research Infrastructure Grant P2C from NICHD: P2C-HD041022, CCPR's Population Research Training Grants T32 from NICHD: T32-HD007545, and the Institute on Global Conflict and Cooperation. The data collection for this project was generously funded by the National Institutes of Health, Population Research Bureau, and the International Initiative for Impact Evaluation. All errors are our own.

3.1 Introduction

Early childhood vaccination and family planning programs are among the most important and widely adopted policies in the developing world over the second half of the twentieth century, yet their long-term effects have been little studied. In particular, these programs' effects on long-run productivity growth through the channel of human capital have rarely been examined.

In this project, we explore long-run effects of a mother and early childhood intervention on productivity in agriculture and enterprises. We focus on a novel channel for health and human capital to affect productivity—technology adoption and firm complexity—and find substantial and positive effects. Finally, the long-term nature of our data allows us to observe program effects decades after implementation, which are substantial and positive, while medium-term effects in agriculture are negligible. While most program evaluations in developing countries only look at short- or medium-run outcomes, doing so in our context would understate the effect of the program we study.

We use uniquely rich data in a subdistrict in Bangladesh to assess the long-run effects of a quasi-experimentally designed maternal and child health and family planning program (MCH-FP). Our data allow us to observe a rich vector of pre-period characteristics as well as medium- and long-run impacts of the program on a variety of relevant enterprise and agricultural outcomes with a large sample size.

Doubt remains about the sign and magnitude of the effect of population health on economic growth. Some argue (e.g., Weil 2007, Bleakley 2010) that health improvements lead to higher income as a result of improved human capital. In contrast, Acemoglu and Johnson (2007) suggest that the productivity-enhancing effects of better health are washed out by increased fertility, a force absent in Weil's 2007 analysis. In our context, we look at the effect of a program which first reduces fertility through policy interventions and then, in a later phase of the intervention, also improves child health. In this setting, the Malthusian channel of increased population competition canceling out the economic benefits of improved health is effectively shut off, and in our results we find an economically significant impact of reduced fertility and improved child health on later-life productivity outcomes. The pairing of a fertility and health program thus provides a unique context to assess the effect of health on productivity and economic growth.

Much of the previous literature has focused on contemporaneous effects of health on agricultural productivity (e.g., Pitt and Rosenzweig 1986 and Fink and Masiye 2015). More generally, Foster and Rosenzweig (1996) find that education raises farm productivity and technology adoption. In addition, Bleakley and Lange (2009) look at the long-run impact of child health improvements on fertility. In contrast, we look at long-term productivity outcomes of a program jointly affecting fertility and child health.

In addition, new management technique and production technology adoption is often slow in the developing world despite the significant financial benefits to farmers and entrepreneurs (Atkin et al., 2017; World Bank, 2007). We hypothesize that an important determinant of new technology and management technique adoption is the human capital of the entrepreneur or farmer, shaped significantly by one's childhood environment (Heckman, 2006), and a channel thus far understudied in the literature.

3.2 Background and Data

We focus on the rural subdistrict of Matlab in Bangladesh as the context of our study. Matlab provides a unique context for two reasons. First, a rich set of data has been collected in the region since 1971, including demographic surveillance information on all inhabitants of Matlab, multiple survey rounds containing various socioeconomic information, and two complete census waves in 1974 and 1982. A second key advantage of studying Matlab is that in 1977 a major maternal and child health and family planning program (MCH-FP) was rolled out to half the subdistrict. The MCH-FP rollout balanced pre-program covariates between treatment and control regions and provides the quasi-experimental variation that we leverage to identify causal effects in this paper (Phillips et al., 1982).

3.2.1 Intervention

In 1977 the MCH-FP program was introduced to half the Matlab subdistrict of Bangladesh by the International Center for the Diarrhoeal Disease Research, Bangladesh (icddr,b), with the remaining half serving as the control. The rollout of the program balanced characteristics between treatment and control regions prior to the implementation of the program and provides the quasi-experimental variation that we leverage to identify causal effects in this paper (Phillips et al., 1982) (see Table 3.1 for a balance test at the household level). Prior research on the MCH-FP has extensively documented the similarity in mortality and fertility rates and household and individual characteristics prior to the intervention (Koenig et al., 1990; Menken and Phillips, 1990; Joshi and Schultz, 2013). Barham (2012) shows that cognition, height, and education were similar after the intervention for individuals too old to be directly affected by the MCH-FP. Barham et al. (2019) also show that pre-program labor market outcomes were balanced as well as just before the child health interventions were rolled out in 1982.

Recent research on the long-run effects of the MCH-FP has found positive impacts. Barham et al. (2019) find that the program led to increased entrepreneurship and occupational upgrading 35 years later for the most affected children. Barham et al. (2021) find that the program resulted in sustained improvements in height.

Program interventions were phased in, with the family planning program and tetanus toxoid vaccines for pregnant women distributed in the first few years (1977 to 1982), a measles vaccine for children starting in 1982, and other vaccines for tetanus, pertussis, polio, and tuberculosis for children distributed starting in 1985. After 1988, the program expanded to the control area. This staggered rollout of program components led to differential treatment of children depending on their year of birth.

The MCH-FP program availability depended on one's location of residence during the program period. Since households may have selectively changed location in response to the program and so their location during the program is potentially endogenous (Barham and Kuhn, 2014), we use Demographic Surveillance System (DSS) and census data to generate an intent-to-treat indicator for each individual based on the pre-program village of residence in 1974. We link individuals to their pre-program household in the following way: (i) link respondent to the 1974 census through the household head of their first residence in the DSS area, (ii) if their household head was absent for the 1974 census, we identify that person's first household head in the DSS area and link that new person to the 1974 census, (iii) we assign remaining unlinked respondents a treatment status using the location of their household head in the DSS area after the 1974 census, but before the start of the MCH-FP in 1977.

3.2.2 Data

We leverage several unusually detailed datasets collected in the Matlab study area. Moreover, we link each of these datasets at the individual or household level, using unique individual and household identifiers provided in the data.

To measure firm and farm outcomes, we use both the 1996 (MHSS1) and 2012 (MHSS2) waves of the Matlab Health and Socioeconomic Survey panel. These data contain a rich set of household agricultural variables, including inputs (e.g., acres, spending on inputs, use of high-yield seeds) and output (quantity harvested, revenue, profits) for 11 types of crops. We also observe a rich set of outcomes for household enterprises. In addition, we use the 1974 and 1982 censuses to obtain baseline characteristics.

To assign treatment status to individuals and households, we use Demographic Surveillance System (DSS) data. These data track every life event—births, deaths, marriages, separations, and migrations—for every inhabitant of Matlab, as well as where each person lives within Matlab, since 1974. The DSS data allow us to observe the family tree of individuals living in Matlab and trace back whether an individual's antecedents lived in the treatment or control area. We use this traceback of antecedent treatment status back in 1974 to assign intent-to-treat status, as we detail below. We exploit the fact that each individual has a unique ID, allowing us to link across DSS, census, and MHSS1 and MHSS2 datasets.

Due to the long-term nature of our outcomes 35 years after the conclusion of the

MCH-FP experimental period, it was crucial to minimize attrition. To accomplish this, the tracking protocol for the data collection followed internal migrants throughout Bangladesh, interviewed international migrants when they returned for holidays or over the phone. This comprehensive tracking was a key feature of the design of this study and had substantial success. Thirty-five years after the start of the MCH-FP, we interviewed over 90 percent of men born during the experimental program—the group with the highest migration rates. Response rates for females and other age groups are even higher. Relative to other studies covering similar populations and longitudinal studies that cover shorter time periods, these rates of attrition are remarkably low.

3.3 Enterprise Results

We estimate the effect of the MCH-FP on enterprise outcomes using variation between treatment and control villages and the timing of the rollout of program components to different villages. We find a significant and positive effect of the intervention on enterprise profits, but effects on other outcomes are sensitive to using sample weights.

We estimate a single-difference equation of the form:

$$Y_{iv} = \beta_0 + \beta_1 T_v + \beta_2 Born_i^{70-77} + \beta_3 Born_i^{77-82} + \beta_4 Born_i^{82-88} + \beta_5 (T_v \times Born_i^{70-77}) + \\ \beta_6 (T_v \times Born_i^{77-82}) + \beta_7 (T_v \times Born_i^{82-88}) + \alpha_{y(i)} + \gamma X_i + \epsilon_{iv} \quad (3.1)$$

for individual i from antecdent village v. $Born_i^{y_1-y_2}$ is an indicator variable for whether i was born between years y_1 and y_2. T_v is an indicator for whether i is treated; $\alpha_{y(i)}$ is a set of indicator variables for i's birth year; and X_i is the vector of demographic and baseline characteristics detailed in Table 3.1. We cluster standard errors by the 1974 village of i (or i's antecedents if i was not born by 1974).

The coefficients β_5, β_6, and β_7 represent the intent-to-treat single-difference coefficients. In particular, they represent the difference in conditional means for the outcome for the relevant age group. Past research on the effects of the MCH-FP by Barham (2012) and Barham et al. (2019) have found pronounced effects for the cohorts born between 1982 and 1988, which is consistent with the findings we present here on enterprises. We focus on men, who are nearly twice as likely to engage in entrepreneurship than women.[2] In addition, we restrict our sample to individuals between the ages of 24 and 65; therefore, the reference group are men born prior to 1970.

[2]We find no statistically significant effect of the MCH-FP on women's enterprise outcomes, as shown in Table 3.15.

Results. Our results are shown in Table 3.3. We find that the program led to substantially higher profits for firms owned and operated by individuals treated with the early-life health interventions.[3] We also find suggestive evidence that treated individuals found more enterprises, are more likely to keep a dedicated account for their business, and employ more workers. However, these latter results do not remain statistically significant when we use sample weights that account for non-random sampling and household combinations, as we show in Tables 3.19, 3.21, and 3.23.

Mechanisms. We find that the positive effect of the MCH-FP on enterprise profits is driven by treated individual's human capital, both physical and cognitive. In particular, as we show in column 2 of Tables 3.9 and 3.11, one's height and MMSE score positively impact firm profits. In contrast, the Malthusian force of the number of competing male siblings does not significantly affect one's own firm profits.

We also leverage the staggered rollout of program components across Matlab. As we detail in Section 3.2.1 and Table 3.12, treatment villages received the child health interventions at different times. In particular, children under the age of 5 in treatment areas A and C (Treatment Area 1; see Figure 3.1) were eligible to receive measles vaccination starting in March 1982, while children in the same age group living in treatment areas B and D (Treatment Area 2) had to wait until November 1985. This variation suggests that looking at the difference in outcomes for individuals born between 1982 and 1984 and living in Treatment Area 1 as compared to those living in Treatment Area 2 may be useful for determining the extent to which childhood vaccinations contribute to the total estimated program effects.

Our estimation results split by treatment area for those born between 1982 and 1984 are shown in Table 3.7. The coefficient size is economically larger for those born in Treatment Area 1, whcih received the measles vaccine, further suggesting the improved early-life human capital resulting from the measles vaccine improved later-life outcomes. However, the effect is not statistically significant.

Weighting. In the above results for enterprise outcomes, we do not weight our estimates. This could lead to inconsistent estimates if sampling was endogenous (Solon et al., 2015). Foster and Milusheva (2017) note that the MHSS in 1996 oversampled individuals who lived in baris with few households in 1974 and increasingly oversampled individuals as their number of antecendent 1974 households increased.[4] We address this oversampling

[3] We also note that the effect of the treatment on the reference group–those born prior to 1970–is negative. This can rationalize our findings of positive effects of the program on the more intensively treated with the findings in Barham et al., 2019 that consumption did not rise as a result of the MCH-FP. In particular, there appears to be some crowding out of enterprises by the younger, more treated entrepreneurs.

[4] Another plausible source of endogenous sampling is attrition between MHSSS1 and MHSS2 surveys. Barham et al. (2019) show attrition was extremely low and fairly balanced acrosss observables between

in the unweighted estimation by controlling for an individual's baseline 1974 bari and family size.

However, because we do not directly control for the variables that determine oversampling, we re-estimate our main enterprise regressions using the inverse sampling probability weights developed by Foster and Milusheva (2017). In particular, we are concerned about either (1) the MHSS2 individual's antecedent's number of households in their bari in 1974, or (2) the number of distinct antecedent households to an individual's 1996 household, being correlated with 2012 enterprise outcomes.

We compare our unweighted and weighted results in Tables 3.19, 3.21, 3.23, and 3.25. We find that for many outcomes, the coefficient of interest (born 1982-1988 interacted with the treatment dummy), loses signifcance and sometimes changes sign. This is the case for the number of businesses owned, shown in Table 3.19. Statistical significance is lost for the outcomes of having a dedicated bank account for one's business and the number of employees, though the sign remains positive. Importantly, the total value of enterprise profits remains positive and statistically significant.

3.4 Agriculture

We next turn to the effects of the MCH-FP on agricultural productivity and technology adoption.

Specification. We consider agricultural outcomes at the household level for two reasons. First, this is the level on which data was collected in the MHSS surveys; in contrast, enterprise outcomes were surveyed at the individual level. Second, agricultural decisions are typically made at the household level in Matlab, whereas enterprises are often developed on an individual level.[5]

We determine treatment at the household level as follows. A household in 1996 (from the first MHSS survey wave) is considered treated if the household head can be linked to a household back in 1974 that was living in the treatment area or else migrated into the treatment area by 1977. An MHSS2 household is assigned to the treatment group if the household head was living in a treated household (as defined above) during the MHSS1 survey period in 1996.

To estimate the effect of the MCH-FP on agricultural outcomes, we use a single-difference intent-to-treat specification:

treatment and control groups.

[5]Among enterprise owners in our data, nearly 92 percent own the business themselves, with only about 3.5 percent jointly owning the business with another household member.

$$Y_{hv} = \omega_0 + \omega_1 T_v + \zeta X_h + \varepsilon_{hv} \tag{3.2}$$

for household h and village v. T_v is an indicator for whether household h is considered treated and X is the vector of demographic and baseline characteristics detailed in Table 3.1. We cluster standard errors by the village of the household head of h or his antecedents in 1974.

Results. We find that the program had a negligible effect on the propensity to farm and profits-per-acre in the medium-term (columns 1 and 2 of Table 3.2). In the long-term (i.e. by 2012) we find that the program moderately increased the propensity to farm, had a substantially positive effect on profits per acre on the order of a 29% rise relative to the mean, and a large effect (43% increase relative to the mean) on the chance of adopting a new, more productive technology—high-yield seeds (columns 3, 4, and 5 of Table 3.2). The small and statistically insignificant effects in the short-run contrast with the economically large and precisely estimated effects in the long-run. One important contribution of our study is highlighting the need to look at long-run outcomes of policy interventions, currently a shortcoming of many program evaluations in the development literature.

Mechanisms. We first examine the heterogeneity by age of the household head. This is based on the notion (graphically depicted in Figure 3.3) that exposure to the MCH-FP varied by age of the household head since heads in prime fertility age groups will be most exposed. We divide up cohorts into a minimally exposed pre-1939 bin, a highly exposed 1940-49 bin, and moderately exposed 1950-59 bin, and a bin for those born 1960 and after, when both treatment and control area heads received the fertility and child health interventions during their prime parenting years. As shown in Table 3.16, we find that younger, more treated household heads are the most likely to adopt new technology (high-yield seeds) and have the highest increase in profits per acre as a result of the program.

We also estimate the effect of the staggered rollout of the measles vaccine across villages in Matlab between 1982 and 1984, as discussed above. We compare household heads (in 2012) between treatment areas 1 and 2, born between 1982 and 1984. We present results in Table 3.4. While individuals receiving measles vaccines near birth (those in Treatment Area 1) experience less negative higher agricultural profits per acre and takeup of high-yield variety seeds, the cohort of household heads born between 1982 through 1984 achieve worse agricultural outcomes than older household heads.

We also explore how cognition, education, and height shape our agricultural results.

We show our results in Table 3.5 . We find that a higher MMSE score, indicating better cognition, leads to higher agricultural profits and increased takeup of modern high-yield variety seeds.[6]

Finally, we look at how sibling competition shapes agricultural outcomes. Specifically, since the MCH-FP induced families to have fewer children, this may have led to greater investment per child for treated families. This higher investment raised the child's human capital, therefore making them more capable of adopting new technologies and being more productive at work. In addition, a key channel for the number of male siblings to shape agricultural profitability is through land inheritance. In Bangladesh, typically each male son receives an equal share of land following the death of the father (van Schendel and Rahman, 1997). Therefore, having fewer brothers means that a son will inherit relatively more land.

We find some evidence of sibling competition weighing down agricultural outcomes, but no evidence for enterprise outcomes. We show results for agriculture in Table 3.5. As the number of male siblings rises for a household head, profit made from crops per acre declines, as shown in columns 2 and 4. Moreover, in Table 3.17 columns 1 and 2 we show that larger landholdings correspond to greater takeup of high-yield variety seeds and higher crop profits per acre.

In contrast, we find no effect of Malthusian forces on enterprise outcomes, as shown in the last 3 rows of Tables 3.13 and 3.14. We rationalize this differential effect of the number of male siblings as being driven by the nature of land inheritance, since additional brothers are direct competition for a son's agricultural land holdings.

Aggregation. We estimate the effect of the MCH-FP on agricultural outcomes above at the household-level. As noted above, we do this because (1) the data on agricultural outcomes was collected at the household-level, and (2) land holdings and decisions for agriculture are held at the household-level. Additionally, we define our treatment variable to be the treatment status of the household head in 2012.

We alternatively estimate the effect of the MCH-FP on agricultural outcomes at the individual-level. We show our results in Table 3.27. We continue to find positive effects of treatment on the use of high-yield variety seeds and profits from crops per acre. The effect is statistically significant for the use of high-yield seeds, as seen in columns 1 and 3. We do not obtain statistically significance for the effect on crop profits per acre except for the cohort born between 1977 and 1982.

[6]The Mini Mental State Exam (MMSE) is a cognitive exam measuring five areas of cognitive functioning: orientation, attention/concentration, registration, recall, and language (Folstein et al., 1975). A culturally appropriate MMSE for Bangladesh was used for the MHSS surveys. For more, see page 255 of Barham (2012).

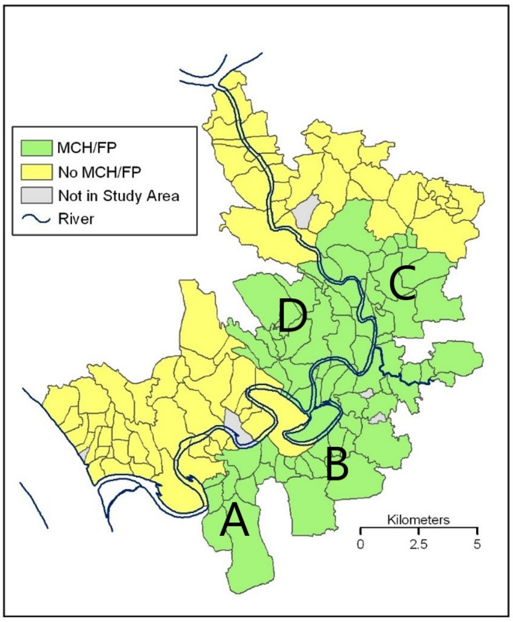
3.5 Conclusion

In this paper, we have provided evidence that a fertility and early-childhood health initiative raises the profitability of firms and farms in the context of rural Bangladesh 35 years after the initiative. We also find that the initiative raised the likelihood of treated households using high-yield variety seeds, and suggestive evidence that firms increased their complexity as a result of the program. These changes were driven by the improvement to individual human capital in the case of firms, and to reduced sibling competition in the case of farms.

We do not find any effect of the initiative 10 years after its conclusion, but do after 35 years, suggesting conventional developing country intervention studies may be limited by short-term assessment periods.

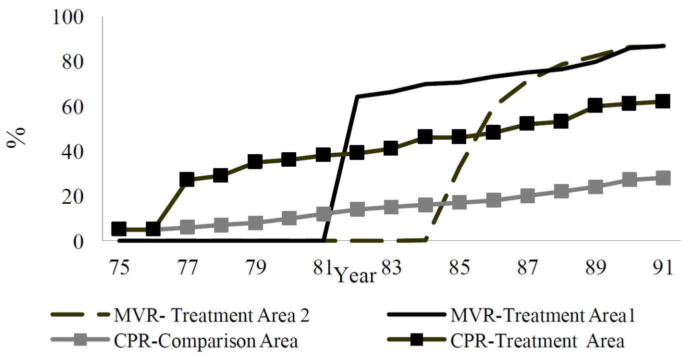
Our results add to the evidence suggesting that improving children's health and providing access to contraception can significantly improve outcomes important for economic development.

Figure 3.1: Map of Matlab Study Area



Notes: Villages within the treatment area (in green) labeled A or C correspond to Treatment Area 1, which began receiving measles vaccines for young children starting in 1982, whereas areas B and D, Treatment Area 2, received measles vaccines as part of the MCH-FP starting in 1985. For more on the rollout, see Table 3.12.

Figure 3.2: Trends in contraceptive prevalence rate (CPR) and measles vaccination rates (MVR) for children 12-59 months by calendar year



Source: Replicated from Figure 2 in Barham et al. (2019).

Table 3.1: Balance at 1974 Baseline

Baseline (1974 Census) Variable	Comparison Area Mean	Comparison Area SD	Treatment Area Mean	Treatment Area SD	Difference in Means Mean	Difference in Means p-value
Land (1982)	11.258	(13.279)	10.807	(18.175)	-0.452	(0.685)
Bari size	8.302	(5.229)	9.628	(6.386)	1.326**	(0.020)
Family size	6.923	(2.867)	6.854	(2.935)	-0.069	(0.743)
Wall tin or tin mix (=1)	0.292	(0.441)	0.301	(0.450)	0.009	(0.770)
Tin roof (=1)	0.814	(0.379)	0.812	(0.384)	-0.002	(0.949)
Number of boats	0.666	(0.586)	0.648	(0.643)	-0.017	(0.769)
Owns a lamp (=1)	0.583	(0.482)	0.621	(0.479)	0.038	(0.371)
Owns a watch (=1)	0.124	(0.320)	0.146	(0.347)	0.022	(0.373)
Owns a radio (=1)	0.067	(0.242)	0.072	(0.254)	0.005	(0.759)
Number of rooms	0.209	(0.097)	0.216	(0.096)	0.007	(0.255)
Number of cows	1.398	(1.748)	1.395	(1.630)	-0.003	(0.982)
Latrine (=1)	0.871	(0.327)	0.781	(0.408)	-0.090**	(0.011)
Drinking water, tubewell (=1)	0.141	(0.339)	0.317	(0.459)	0.176***	(0.000)
Drinking water, tank (=1)	0.351	(0.467)	0.388	(0.480)	0.037	(0.519)
HH head years of education	1.945	(2.869)	2.261	(2.906)	0.316	(0.102)
HH head works in agriculture (=1)	0.586	(0.481)	0.613	(0.480)	0.028	(0.494)
HH head works in fishing (=1)	0.063	(0.237)	0.067	(0.248)	0.004	(0.860)
HH head age	46.587	(13.181)	47.106	(14.034)	0.519	(0.580)
HH head spouse's years of education	0.540	(1.256)	0.687	(1.462)	0.146	(0.105)
HH head spouse's age	36.229	(9.976)	36.738	(10.914)	0.509	(0.486)
Observations	1,302		1,208		2,510	

Notes: Sample includes households observed both in MHSS1 and MHSS2 survey waves. Characteristics refer to baseline in reference to household head (or closest person to head with nonmissing data) at time of MHSS2. Standard deviations are clustered at the 1974 village level. Statistics are weighted by Foster and Milusheva (2017) weights.

Table 3.2: Effects of MCHFP: Agriculture

	MHSS1 (1996)		MHSS2 (2012-2014)		
	(1)	(2)	(3)	(4)	(5)
	=1 if household farms	Agr profit per acre	=1 if household farms	Agr profit per acre	=1 if uses high-yield seed
Treat	0.0471	-14.99	0.0547*	83.49**	0.0901**
	(0.0442)	(13.14)	(0.0316)	(36.63)	(0.0350)
Observations	1836	1836	2510	2510	2510
Adjusted R^2	0.054	0.038	0.086	0.037	0.027
Baseline controls	Y	Y	Y	Y	Y
% chg. rel. to mean	7.2	-15.7	8.5	29.4	43.4

Notes: This table shows our estimates of $Y_{hv} = \omega_0 + \omega_1 Treat_v + \zeta BaselineControls_h + \varepsilon_{hv}$ for MHSS1 (columns 1 and 2) and MHSS2 (columns 3-5). Standard errors clustered at the baseline village level. Weights from Foster and Milusheva (2017). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.3: Long-run Effects of MCHFP: Enterprises (Men)

	(1) Num. business owned	(2) Have business loans	(3) Keep dedicated bank acct. for business	(4) Keep detailed accts. for business
Treat	-0.0175	0.0409*	-0.0227**	-0.0236
	(0.0520)	(0.0223)	(0.00978)	(0.0251)
Treat × Born 1970-1977	-0.101	-0.0915**	0.0384*	0.0322
	(0.0827)	(0.0367)	(0.0206)	(0.0336)
Treat × Born 1977-1982	0.137*	-0.0407	0.0141	0.00606
	(0.0763)	(0.0355)	(0.0195)	(0.0446)
Treat × Born 1982-1988	0.120*	-0.0202	0.0348**	0.0444
	(0.0651)	(0.0249)	(0.0173)	(0.0347)
Observations	2911	2911	2911	2911
Mean if born 1970-77	0.69	0.16	0.03	0.17
Mean if born 1977-82	0.44	0.10	0.04	0.16
Mean if born 1982-88	0.33	0.04	0.03	0.10
	(1) Num. workers employed	(2) Tot. profits (taka)	(3) Revenue	(4) Revenue per worker
Treat	-0.480*	-19760.3**	3733.3	61.84
	(0.282)	(9413.3)	(6559.8)	(1741.2)
Treat × Born 1970-1977	0.119	43978.3	1109.9	1719.5
	(0.333)	(36819.1)	(10111.8)	(4305.4)
Treat × Born 1977-1982	0.417	37654.1	-32512.6**	-10106.4**
	(0.317)	(38433.8)	(13850.5)	(4567.1)
Treat × Born 1982-1988	0.598**	102089.7**	4014.9	21.96
	(0.283)	(50779.2)	(12200.3)	(3705.2)
Observations	2911	2911	2911	2911
Mean if born 1970-77	1.2	69036.1	17456.3	10466.0
Mean if born 1977-82	0.8	56948.6	37200.9	15115.8
Mean if born 1982-88	0.5	33017.1	14608.2	6846.5

Notes: The table presents estimates of $Y_{iv} = \beta_0 + \beta_1 T_v + \beta_2 Born_i^{70-77} + \beta_3 Born_i^{77-82} + \beta_4 Born_i^{82-88} + \beta_5(T_v \times Born_i^{70-77}) + \beta_6(T_v \times Born_i^{77-82}) + \beta_7(T_v \times Born_i^{82-88}) + \alpha_{y(i)} + \gamma X_i + \epsilon_{iv}$ at the individual-level, where T_v is a treatment indicator at the 1974 village level, X_i is a vector of baseline controls, and $\alpha_{y(i)}$ are year of birth fixed effects. Means by age group are for the comparison group. Standard errors are clustered by pre-program village. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.4: Agriculture Mechanisms: Treatment Area

	(1)	(2)
	Agr profit/acre	=1 if used high-yield seed
Treat	79.92*	0.0894**
	(42.27)	(0.0371)
Treat × Head born 1977-81	120.6	0.0374
	(103.8)	(0.0770)
Treat × Head born 1985-88	-36.82	0.0195
	(79.07)	(0.0687)
Treat × Head born 1989+	122.8	0.0231
	(81.98)	(0.0730)
Treat Area 1 × Head born 1982-84	-137.5	-0.0272
	(150.4)	(0.227)
Treat Area 2 × Head born 1982-84	-413.9***	-0.297***
	(100.9)	(0.0906)
Head born 1982-84	-3.517	-0.00928
	(74.98)	(0.0793)
Head born 1985-88	-232.2***	-0.164***
	(57.66)	(0.0378)
Head born 1977-81	-134.7***	-0.0860*
	(43.67)	(0.0454)
Head born 1989+	-361.1***	-0.230***
	(71.33)	(0.0456)
Observations	2510	2510
R^2	0.059	0.043
1974 baseline controls	Y	Y
Weighted	Y	Y

Notes: Standard errors are clustered by pre-program village. Sample weights from Foster and Milusheva (2017). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.5: Agriculture Mechanisms: Human Capital

	(1)	(2)	(3)	(4)
	Agr profit/acre	=1 if used high-yield seed	Agr profit/acre	=1 if used high-yield seed
Treat	92.18**	0.0897**	183.9***	0.168**
	(41.49)	(0.0401)	(63.21)	(0.0712)
Num. of brothers of HH head	-22.27***	-0.00433	-6.525	0.00911
	(8.269)	(0.00874)	(10.53)	(0.00799)
Highest grade attended (Z-score for those born 1982-88)	0.336	-0.0102	-18.92	-0.0183
	(27.54)	(0.0305)	(28.94)	(0.0340)
MMSE score (Z-score for those born 1982-88)	-7.372	0.0377**	32.79	0.0483**
	(26.82)	(0.0184)	(35.41)	(0.0219)
Height (Z-score for those born 1982-88)	-10.55	-0.0107	-9.212	0.0268
	(39.50)	(0.0321)	(44.18)	(0.0408)
Treat × Highest grade attended (Z-score for those born 1982-88)			27.59	0.0110
			(52.70)	(0.0608)
Treat × MMSE score (Z-score for those born 1982-88)			-57.73	-0.00851
			(44.16)	(0.0329)
Treat × Height (Z-score for those born 1982-88)			4.387	-0.0607
			(74.44)	(0.0621)
Treat × Num. of brothers of HH head			-34.63**	-0.0300*
			(15.16)	(0.0175)
Observations	1843	1843	1843	1843
R^2	0.060	0.040	0.064	0.045
1974 baseline controls	Y	Y	Y	Y
Weighted	Y	Y	Y	Y
Dep. var. mean	150.74	0.11	150.74	0.11

Notes: Standard errors are clustered by pre-program village. Sample weights from Foster and Milusheva (2017). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.6: Enterprise Mechanisms: Treatment Area

	(1) Num. business owned	(2) Have business loans	(3) Keep dedicated bank acct. for business	(4) Keep detailed accts. for business
Treat	0.0739	0.0302	-0.0183*	0.0147
	(0.0693)	(0.0326)	(0.00974)	(0.0389)
Treat × Born 1970-1977	-0.158	-0.0154	0.0207	-0.000649
	(0.118)	(0.0592)	(0.0244)	(0.0560)
Treat × Born 1977-1982	0.0568	-0.000755	0.0386*	-0.0269
	(0.102)	(0.0430)	(0.0213)	(0.0643)
Treat Area 1 × Born 1982-1984	0.00314	-0.0759**	0.00837	-0.0459
	(0.112)	(0.0342)	(0.0230)	(0.0529)
Treat Area 2 × Born 1982-1984	0.107	0.0538	0.0141	0.0598
	(0.121)	(0.0636)	(0.0177)	(0.0720)
Treat × Born 1985-1988	-0.0695	-0.0404	0.0395	0.0335
	(0.101)	(0.0414)	(0.0271)	(0.0618)
Observations	2873	2873	2873	2873
Mean if born 1970-77	0.69	0.16	0.03	0.17
Mean if born 1977-82	0.44	0.10	0.04	0.16
Mean if born 1982-88	0.33	0.04	0.03	0.10

Notes: The table presents estimates at the individual-level. Standard errors are clustered by pre-program village. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.7: Enterprise Mechanisms: Treatment Area

	(1) Num. workers employed	(2) Tot. profits (taka)	(3) Revenue	(4) Revenue per worker
Treat	-0.0169	-20529.0**	2479.5	-900.4
	(0.165)	(9678.5)	(4343.0)	(2283.1)
Treat × Born 1970-1977	-0.150	17344.2	-1862.8	124.2
	(0.262)	(16966.9)	(7400.5)	(4124.6)
Treat × Born 1977-1982	-0.128	-4309.9	-26250.0**	-6968.3
	(0.286)	(28334.9)	(10508.2)	(4596.9)
Treat Area 1 × Born 1982-1984	-0.0945	90617.4	-3695.4	-220.5
	(0.299)	(59208.6)	(13446.6)	(5344.3)
Treat Area 2 × Born 1982-1984	0.157	12883.6	-12434.2	-171.0
	(0.322)	(21900.6)	(8330.9)	(2894.3)
Treat × Born 1985-1988	-0.0141	59193.6	-9731.2	-3600.3
	(0.243)	(43279.6)	(12542.6)	(6307.5)
Observations	2873	2873	2873	2873
Mean if born 1970-77	1.20	69036.09	17456.32	10466.02
Mean if born 1977-82	0.85	56948.57	37200.93	15115.83
Mean if born 1982-88	0.48	33017.14	14608.23	6846.50

Notes: The table presents estimates at the individual-level. Standard errors are clustered by pre-program village. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.8: Enterprise Mechanisms: Human Capital

	(1) Num. business owned	(2) Have business loans	(3) Keep dedicated bank acct. for business	(4) Keep detailed accts. for business
Treat	0.0429	0.0537	-0.0144	0.0144
	(0.0755)	(0.0363)	(0.0101)	(0.0371)
Treat × Born 1970-1977	-0.201	0.0489	0.0386	0.0730
	(0.147)	(0.0902)	(0.0370)	(0.0892)
Treat × Born 1977-1982	0.198*	-0.0249	0.0287	-0.00816
	(0.118)	(0.0548)	(0.0262)	(0.0824)
Treat × Born 1982-1988	0.0545	-0.0383	0.0118	0.0296
	(0.114)	(0.0463)	(0.0253)	(0.0570)
M2 MMSE Score	-0.00446	0.00463**	0.00206**	0.0135***
	(0.00510)	(0.00232)	(0.000824)	(0.00249)
M2 height	-0.283	-0.0454	0.141	-0.0996
	(0.328)	(0.164)	(0.0969)	(0.181)
Number of brothers	-0.00391	0.00189	0.00153	-0.00735
	(0.0138)	(0.00724)	(0.00416)	(0.00756)
Observations	1898	1898	1898	1898
Mean if born 1970-77	0.69	0.16	0.03	0.17
Mean if born 1977-82	0.44	0.10	0.04	0.16
Mean if born 1982-88	0.33	0.04	0.03	0.10

Notes: The table presents estimates at the individual-level. Standard errors are clustered by pre-program village. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.9: Enterprise Mechanisms: Human Capital

	(1) Num. workers employed	(2) Tot. profits (taka)	(3) Revenue	(4) Revenue per worker
Treat	0.0192	-29714.4**	6920.7	559.7
	(0.186)	(13919.8)	(5134.1)	(1775.4)
Treat × Born 1970-1977	-0.121	14107.4	1682.7	2963.6
	(0.326)	(29218.8)	(11292.6)	(7214.7)
Treat × Born 1977-1982	-0.381	-23869.4	-33717.5**	-5919.3
	(0.400)	(46990.6)	(13643.2)	(4907.7)
Treat × Born 1982-1988	0.0625	91529.9	-18009.0	-5899.8
	(0.318)	(62962.2)	(14456.6)	(7074.3)
M2 MMSE Score	0.0226*	1758.3	925.4*	510.9**
	(0.0130)	(1135.9)	(492.2)	(208.5)
M2 height	1.666	308301.7	70599.2*	22777.7*
	(1.284)	(236448.8)	(37544.0)	(12675.5)
Number of brothers	0.00467	584.0	1806.6	989.4
	(0.0561)	(7659.2)	(1814.8)	(669.4)
Observations	1898	1898	1898	1898
Mean if born 1970-77	1.20	69036.09	17456.32	10466.02
Mean if born 1977-82	0.85	56948.57	37200.93	15115.83
Mean if born 1982-88	0.48	33017.14	14608.23	6846.50

Notes: The table presents estimates at the individual-level. Standard errors are clustered by pre-program village. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.10: Enterprise Mechanisms: Human Capital

	(1) Num. business owned	(2) Have business loans	(3) Keep dedicated bank acct. for business	(4) Keep detailed accts. for business
Treat	0.623	-0.700	-0.0238	-0.653
	(1.082)	(0.513)	(0.256)	(0.619)
Treat × Born 1970-1977	-0.148	0.0431	0.0424	0.0822
	(0.151)	(0.0933)	(0.0369)	(0.0926)
Treat × Born 1977-1982	0.269**	-0.0367	0.0331	-0.00389
	(0.126)	(0.0554)	(0.0289)	(0.0860)
Treat × Born 1982-1988	0.134	-0.0521	0.0166	0.0335
	(0.126)	(0.0477)	(0.0256)	(0.0616)
M2 MMSE Score	0.000812	0.00214	0.00214**	0.0110***
	(0.00520)	(0.00231)	(0.00104)	(0.00246)
M2 height	-0.206	-0.285	0.131	-0.329
	(0.486)	(0.216)	(0.121)	(0.268)
Number of brothers	0.0157	0.00624	0.00368	0.00524
	(0.0156)	(0.00776)	(0.00555)	(0.00885)
Treat × M2 MMSE Score	-0.0114	0.00534	-0.000140	0.00561
	(0.00991)	(0.00500)	(0.00183)	(0.00463)
Treat × M2 height	-0.0852	0.386	0.0160	0.357
	(0.692)	(0.325)	(0.176)	(0.389)
Treat × Number of brothers	-0.0457*	-0.0102	-0.00505	-0.0296*
	(0.0262)	(0.0172)	(0.00835)	(0.0157)
Observations	1898	1898	1898	1898

Notes: The table presents estimates at the individual-level. Standard errors are clustered by pre-program village. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.11: Enterprise Mechanisms: Human Capital

	(1) Num. workers employed	(2) Tot. profits (taka)	(3) Revenue	(4) Revenue per worker
Treat	-8.676**	-813448.1	-44799.6	8748.1
	(3.482)	(676129.3)	(112142.0)	(37162.0)
Treat × Born 1970-1977	-0.132	28328.3	475.9	2339.8
	(0.362)	(31966.1)	(11567.6)	(6793.0)
Treat × Born 1977-1982	-0.410	-608.3	-35484.8**	-6757.9
	(0.455)	(47459.7)	(14679.0)	(5695.3)
Treat × Born 1982-1988	0.0232	116744.9	-20008.2	-6820.2
	(0.337)	(74599.1)	(15985.8)	(8437.3)
M2 MMSE Score	0.00995	3835.5**	728.9	462.4
	(0.0170)	(1482.3)	(758.5)	(296.0)
M2 height	-1.369	-24282.5	55581.9	27239.3*
	(1.250)	(120996.4)	(63994.9)	(15923.1)
Number of brothers	0.0636	9950.7*	1678.5	657.8
	(0.0857)	(5892.7)	(2042.3)	(1080.0)
Treat × M2 MMSE Score	0.0250	-5010.7	418.8	110.0
	(0.0205)	(3116.8)	(905.6)	(364.8)
Treat × M2 height	5.140**	607846.8	24040.2	-8211.6
	(2.232)	(479645.0)	(72269.5)	(22351.1)
Treat × Number of brothers	-0.136	-21450.6	301.1	771.1
	(0.109)	(16171.7)	(3400.7)	(1777.6)
Observations	1898	1898	1898	1898

Notes: The table presents estimates at the individual-level. Standard errors are clustered by pre-program village. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Appendix

Table 3.12: MCH-FP Interventions by Cohort

Birth year	Age in 2012	Program Eligibility
Jan. 1947–Sept. 1977	35–65	No effect during early childhood except indirectly, e.g., through sibling competition.
Oct. 1977–Feb. 1982	31–34	Family planning and maternal health interventions: mothers eligible for family planning, tetanus toxoid vaccine, and folic acid and iron in last trimester of pregnancy.
March 1982–Dec. 1988	24–30	Child health interventions added
March 1982–Oct. 1985	27–30	Interventions added in Treatment Areas A and C: children under age five eligible for measles vaccination
Nov. 1985–Dec. 1988	24–26	Interventions extended to entire treatment area: Children under age five eligible for all vaccines (measles, DPT, polio, tuberculosis), vitamin A supplementation, and nutrition rehabilitation for children at risk starting in 1987.

Notes: This table is based on Table 1 of Barham (2012) and Table A1 of Barham et al. (2019).

Table 3.13: Enterprise Mechanisms: Human Capital

	(1) Num. business owned	(2) Have business Loans	(3) Keep dedicated bank acct. for business	(4) Keep detailed accts. for business
Treat	0.0417	0.0532	-0.0146	0.0195
	(0.0748)	(0.0361)	(0.0101)	(0.0397)
Treat × Born 1970-1977	-0.200	0.0214	0.0344	0.0660
	(0.139)	(0.0839)	(0.0368)	(0.0866)
Treat × Born 1977-1982	0.191	-0.0231	0.0202	-0.00828
	(0.117)	(0.0541)	(0.0243)	(0.0813)
Treat × Born 1982-1988	0.0579	-0.0286	0.0166	0.0335
	(0.115)	(0.0461)	(0.0232)	(0.0571)
Born 1970-1977 × M2 MMSE Score	-0.0265*	-0.0142*	0.00147	0.0000797
	(0.0138)	(0.00825)	(0.00268)	(0.00718)
Born 1977-1982 × M2 MMSE Score	-0.0248**	0.00485	-0.000332	0.0208***
	(0.0114)	(0.00536)	(0.00286)	(0.00713)
Born 1982-1988 × M2 MMSE Score	-0.00749	0.00124	0.000985	0.00449
	(0.00810)	(0.00376)	(0.00172)	(0.00426)
Born 1970-1977 × M2 height	0.163	-0.929	0.109	-0.551
	(0.951)	(0.653)	(0.155)	(0.631)
Born 1977-1982 × M2 height	-0.576	0.110	0.304	-0.826
	(0.831)	(0.379)	(0.190)	(0.503)
Born 1982-1988 × M2 height	-0.637	0.0923	0.102	-0.216
	(0.662)	(0.226)	(0.292)	(0.344)
Born 1970-1977 × Number of brothers	0.0272	-0.0458**	-0.00824	0.00478
	(0.0425)	(0.0194)	(0.00724)	(0.0263)
Born 1977-1982 × Number of brothers	0.00172	0.0113	-0.00786	0.00456
	(0.0331)	(0.0143)	(0.00656)	(0.0184)
Born 1982-1988 × Number of brothers	-0.00670	0.0151	0.00785	-0.00217
	(0.0276)	(0.00949)	(0.00994)	(0.0158)
Observations	1898	1898	1898	1898
Mean if born 1970-77	0.69	0.16	0.03	0.17
Mean if born 1977-82	0.44	0.10	0.04	0.16
Mean if born 1982-88	0.33	0.04	0.03	0.10

Notes: The table presents estimates at the individual-level. Standard errors are clustered by pre-program village. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Figure 3.3: Fertility in Matlab by Birth Year

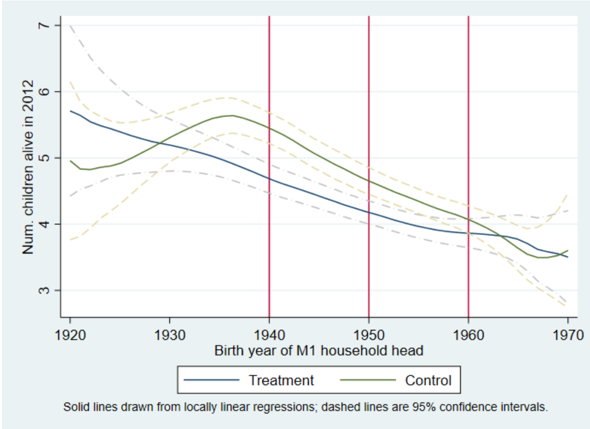


Table 3.14: Enterprise Mechanisms: Human Capital

	(1) Num. workers employed	(2) Tot. profits (taka)	(3) Revenue	(4) Revenue per worker
Treat	0.0220	-27751.0**	6555.1	373.3
	(0.191)	(13280.8)	(5139.2)	(1749.4)
Treat × Born 1970-1977	-0.221	18108.2	4131.2	5629.3
	(0.320)	(29733.2)	(11113.6)	(6513.3)
Treat × Born 1977-1982	-0.461	-10912.3	-33388.8**	-4840.4
	(0.461)	(44980.3)	(14122.0)	(4818.6)
Treat × Born 1982-1988	0.127	85535.5*	-17375.9	-6268.1
	(0.300)	(51318.6)	(13992.0)	(7571.0)
Born 1970-1977 × M2 MMSE Score	-0.0361	2054.5	354.4	772.2
	(0.0292)	(2101.3)	(785.8)	(565.6)
Born 1977-1982 × M2 MMSE Score	0.00502	7710.7*	-543.3	277.1
	(0.0334)	(4439.9)	(2867.8)	(1002.8)
Born 1982-1988 × M2 MMSE Score	-0.0111	-4946.3	154.0	156.9
	(0.0188)	(3686.5)	(553.5)	(352.9)
Born 1970-1977 × M2 height	-1.195	-68324.3	-32007.5	-15669.0
	(2.486)	(140155.4)	(54568.4)	(41436.1)
Born 1977-1982 × M2 height	-1.002	-199671.9	65842.4	29297.2
	(2.308)	(290890.1)	(98950.7)	(38480.9)
Born 1982-1988 × M2 height	-0.166	999084.3	82750.4	24690.8
	(1.350)	(889924.4)	(108512.6)	(27620.0)
Born 1970-1977 × Number of brothers	-0.165	16579.2**	9698.6*	8061.4
	(0.133)	(7443.2)	(5278.0)	(4976.8)
Born 1977-1982 × Number of brothers	-0.0645	16949.3	3480.2	2681.2*
	(0.152)	(12277.3)	(3940.1)	(1529.8)
Born 1982-1988 × Number of brothers	0.0655	-10781.1	1156.3	-543.3
	(0.0817)	(22981.7)	(3588.9)	(1157.2)
Observations	1898	1898	1898	1898
Mean if born 1970-77	1.20	69036.09	17456.32	10466.02
Mean if born 1977-82	0.85	56948.57	37200.93	15115.83
Mean if born 1982-88	0.48	33017.14	14608.23	6846.50

Notes: The table presents estimates at the individual-level. Standard errors are clustered by pre-program village. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.15: Long-run Effects of MCHFP: Enterprises (Women)

	(1)	(2)	(3)	(4)
	Num. business owned	Have business Loans	Keep dedicated bank acct. for business	Keep detailed accts. for business
PANEL A				
Treat	-0.00621	-0.00413	-0.00334	-0.00964*
	(0.0322)	(0.00294)	(0.00306)	(0.00512)
Treat × Born 1970-1977	0.0547	0.00377	0.00793	0.0133
	(0.0562)	(0.00569)	(0.00653)	(0.0115)
Treat × Born 1977-1982	0.0676	0.00232	-0.000600	-0.00410
	(0.0611)	(0.00737)	(0.00461)	(0.00943)
Treat × Born 1982-1988	0.0526	0.00350	-0.00403	0.00595
	(0.0501)	(0.00255)	(0.00501)	(0.00616)
Observations	3396	3396	3396	3396
Mean if born 1970-77	0.42	0.00	0.00	0.02
Mean if born 1977-82	0.24	0.00	0.00	0.01
Mean if born 1982-88	0.21	0.00	0.01	0.01
	(1)	(2)	(3)	(4)
	Num. workers employed	Tot. profits (taka)	Revenue	Revenue per worker
PANEL B				
Treat	-0.00765	24.12	35.42	24.23
	(0.0510)	(302.3)	(48.10)	(44.26)
Treat × Born 1970-1977	-0.0191	-2161.1**	-145.6	-51.44
	(0.0845)	(952.0)	(95.84)	(72.80)
Treat × Born 1977-1982	0.0557	-1114.0	-437.5*	-340.7
	(0.0712)	(791.7)	(261.6)	(206.1)
Treat × Born 1982-1988	0.0552	-360.6	-37.24	-21.64
	(0.0772)	(1301.0)	(116.5)	(117.0)
Observations	3396	3396	3396	3396
Mean if born 1970-77	0.4	3245.0	218.1	119.6
Mean if born 1977-82	0.2	1845.8	425.1	358.9
Mean if born 1982-88	0.2	1828.0	163.2	147.6

Notes: The table presents estimates of equation 3.1 at the individual-level. Means by age group are for the comparison group. Standard errors are clustered by pre-program village. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.16: Agriculture Mechanisms: Parent Age

	(1) =1 if farmer 2012	(2) =1 if used high-yield seed 2012	(3) Agr profit per acre 2012
M1 HH head born pre-1939 X Treat	0.00895	-0.00432	122.6
	(0.0546)	(0.0726)	(83.11)
M1 HH head born 1940-49 X Treat	0.167***	0.104*	107.5
	(0.0572)	(0.0576)	(67.51)
M1 HH head born 1950-59 X Treat	0.0354	0.109**	40.61
	(0.0532)	(0.0462)	(47.91)
M1 HH head born 1960+ X Treat	0.0200	0.113**	80.82*
	(0.0514)	(0.0495)	(48.51)
Observations	2510	2510	2510
Adjusted R^2	0.096	0.028	0.037
Baseline controls	Y	Y	Y
% chg. rel. to mean for born pre-1939	1.3	-2.4	47.9
% chg. rel. to mean for born 1940-49	27.1	52.4	40.4
% chg. rel. to mean for born 1950-59	5.3	49.8	13.5
% chg. rel. to mean for born 1960+	3.2	47.1	25.1

Notes: Standard errors are clustered by pre-program village. Sample weights from Foster and Milusheva (2017). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.17: Agriculture Mechanisms: Land Owned & Emigration

	(1)	(2)	(3)	(4)
	Crop profit per acre	Adopted HYV Seeds	Crop profit per acre	Adopted HYV Seeds
Treat	78.72	0.117**	144.7	0.202
	(39.87)	(0.0369)	(123.3)	(0.118)
Acres owned	33.22*	0.0596***		
	(14.84)	(0.0155)		
Treat × Acres owned	23.19	-0.0391		
	(30.03)	(0.0239)		
% village emigrated			-48.76	0.0446
			(174.7)	(0.164)
Treat × % village emigrated			-193.7	-0.341
			(346.1)	(0.309)
Observations	2470	2470	2510	2510

Notes: Standard errors are clustered by pre-program village. Sample weights from Foster and Milusheva (2017). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.18: Agriculture Mechanisms: Age of HH Members

	(1) =1 if used high-yield seed	(2) =1 if used high-yield seed	(3) Agr profit/acre	(4) Agr profit/acre
Treat	0.0927**		76.74	
	(0.0405)		(47.98)	
Treat=1 × Born 1982-88=1	-0.0117		-8.874	
	(0.0669)		(74.08)	
Treat=1 × Born 1977-81=1	-0.0471		53.49	
	(0.113)		(124.6)	
Treat=1 × Born 1989-96=1	0.0662		149.9	
	(0.107)		(116.8)	
Treat=1 × Born 1972-76=1	-0.0163		-45.57	
	(0.0695)		(82.58)	
% HH treated		0.0152		13.28
		(0.0457)		(51.03)
% HH treated × Frac. HH born 82-88		0.0947		-8.120
		(0.165)		(181.4)
% HH treated × Frac. HH born 77-81		0.100		517.9**
		(0.230)		(247.8)
% HH treated × Frac. HH born 89-96		0.338*		163.6
		(0.179)		(162.0)
% HH treated × Frac. HH born 72-76		0.210		196.6
		(0.195)		(221.3)
Observations	2510	2510	2510	2510
R^2	0.040	0.046	0.055	0.061
1974 baseline controls	Y	Y	Y	Y
Weighted	Y	Y	Y	Y

Notes: Standard errors are clustered by pre-program village. Sample weights from Foster and Milusheva (2017). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.19: Robustness to Weighting: Number of Enterprises Founded

	(1)	(2)	(3)	(4)
	Num. bus. owned	Num. bus. owned	Num. bus. owned	Num. bus. owned
Treat	-0.0175	-0.00641	0.0749	0.0687
	(0.0520)	(0.0508)	(0.0693)	(0.0698)
Treat × Born 1970-1977	-0.101	-0.115	-0.158	-0.167
	(0.0827)	(0.0837)	(0.118)	(0.119)
Treat × Born 1977-1982	0.137*	0.117	0.0562	0.0573
	(0.0763)	(0.0772)	(0.102)	(0.103)
Treat × Born 1982-1988	0.120*	0.108*	-0.0156	-0.0128
	(0.0651)	(0.0640)	(0.0917)	(0.0917)
Observations	2911	2873	2873	2873
Weights	Unweighted	Unweighted	Foster-Milusheva	Backbone weight
Sample	men	Weights nonmissing	Weights nonmissing	Weights nonmissing
Controls	all	all	all	drop 74 bari/fam size

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.20: Robustness to Weighting: Have Business Loans

	(1)	(2)	(3)	(4)
	Any business loans	Any business loans	Any business loans	Any business loans
Treat	0.0409*	0.0420*	0.0311	0.0312
	(0.0223)	(0.0227)	(0.0325)	(0.0322)
Treat × Born 1970-1977	-0.0915**	-0.101***	-0.0158	-0.0149
	(0.0367)	(0.0372)	(0.0590)	(0.0597)
Treat × Born 1977-1982	-0.0407	-0.0450	-0.00107	-0.000946
	(0.0355)	(0.0353)	(0.0431)	(0.0429)
Treat × Born 1982-1988	-0.0202	-0.0211	-0.0246	-0.0247
	(0.0249)	(0.0252)	(0.0388)	(0.0386)
Observations	2911	2873	2873	2873
Weights	Unweighted	Unweighted	Foster-Milusheva	Backbone weight
Sample	men	Weights nonmissing	Weights nonmissing	Weights nonmissing
Controls	all	all	all	drop 74 bari/fam size

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.21: Robustness to Weighting: Have Dedicated Business Bank Account

	(1)	(2)	(3)	(4)
	Dedicated bank acct.	Dedicated bank acct.	Dedicated bank acct.	Dedicated bank acct.
Treat	-0.0227**	-0.0227**	-0.0183*	-0.0177*
	(0.00978)	(0.00986)	(0.00976)	(0.00981)
Treat × Born 1970-1977	0.0384*	0.0354*	0.0206	0.0224
	(0.0206)	(0.0208)	(0.0243)	(0.0237)
Treat × Born 1977-1982	0.0141	0.0150	0.0387*	0.0388*
	(0.0195)	(0.0196)	(0.0213)	(0.0213)
Treat × Born 1982-1988	0.0348**	0.0356**	0.0281	0.0278
	(0.0173)	(0.0174)	(0.0190)	(0.0190)
Observations	2911	2873	2873	2873
Weights	Unweighted	Unweighted	Foster-Milusheva	Backbone weight
Sample	men	Weights nonmissing	Weights nonmissing	Weights nonmissing
Controls	all	all	all	drop 74 bari/fam size

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.22: Robustness to Weighting: Keep Detailed Accounts

	(1) Detailed accts.	(2) Detailed accts.	(3) Detailed accts.	(4) Detailed accts.
Treat	-0.0236	-0.0226	0.0153	0.0130
	(0.0251)	(0.0255)	(0.0390)	(0.0390)
Treat × Born 1970-1977	0.0322	0.0257	-0.00120	-0.00306
	(0.0336)	(0.0335)	(0.0559)	(0.0572)
Treat × Born 1977-1982	0.00606	-0.00609	-0.0270	-0.0263
	(0.0446)	(0.0451)	(0.0644)	(0.0646)
Treat × Born 1982-1988	0.0444	0.0431	0.0258	0.0266
	(0.0347)	(0.0349)	(0.0524)	(0.0522)
Observations	2911	2873	2873	2873
Weights	Unweighted	Unweighted	Foster-Milusheva	Backbone weight
Sample	men	Weights nonmissing	Weights nonmissing	Weights nonmissing
Controls	all	all	all	drop 74 bari/fam size

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.23: Robustness to Weighting: Employees

	(1)	(2)	(3)	(4)
	num. employees	num. employees	num. employees	num. employees
Treat	-0.480*	-0.473*	-0.0152	-0.0232
	(0.282)	(0.284)	(0.165)	(0.165)
Treat × Born 1970-1977	0.119	0.108	-0.151	-0.169
	(0.333)	(0.339)	(0.262)	(0.261)
Treat × Born 1977-1982	0.417	0.391	-0.129	-0.128
	(0.317)	(0.322)	(0.286)	(0.285)
Treat × Born 1982-1988	0.598**	0.590**	0.0119	0.0162
	(0.283)	(0.287)	(0.237)	(0.237)
Observations	2911	2873	2873	2873
Weights	Unweighted	Unweighted	Foster-Milusheva	Backbone weight
Sample	men	Weights nonmissing	Weights nonmissing	Weights nonmissing
Controls	all	all	all	drop 74 bari/fam size

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.24: Robustness to Weighting: Total Profit

	(1)	(2)	(3)	(4)
	Tot. profits (30 day)	Tot. profits (30 day)	Tot. profits (30 day)	Tot. profits (30 day)
Treat	-19760.3**	-19537.1**	-21020.9**	-21197.4**
	(9413.3)	(9569.3)	(9668.4)	(9766.5)
Treat × Born 1970-1977	43978.3	45890.8	17628.5	16947.7
	(36819.1)	(37572.5)	(16943.4)	(16624.7)
Treat × Born 1977-1982	37654.1	3373.3	-4156.3	-4197.8
	(38433.8)	(19317.8)	(28334.2)	(28382.0)
Treat × Born 1982-1988	102089.7**	104084.2**	53851.8*	53976.2*
	(50779.2)	(51096.2)	(31553.9)	(31737.7)
Observations	2911	2873	2873	2873
Weights	Unweighted	Unweighted	Foster-Milusheva	Backbone weight
Sample	men	Weights nonmissing	Weights nonmissing	Weights nonmissing
Controls	all	all	all	drop 74 bari/fam size

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.25: Robustness to Weighting: Total Revenue

	(1)	(2)	(3)	(4)
	Total revenue (30 day)	Total revenue (30 day)	Total revenue (30 day)	Total revenue (30 day)
Treat	3733.3	4124.5	2431.1	2669.9
	(6559.8)	(6530.5)	(4335.5)	(4371.9)
Treat × Born 1970-1977	1109.9	499.6	-1819.6	-779.0
	(10111.8)	(10143.6)	(7415.0)	(7211.2)
Treat × Born 1977-1982	-32512.6**	-33547.3**	-26242.0**	-26165.2**
	(13850.5)	(13871.2)	(10494.4)	(10525.9)
Treat × Born 1982-1988	4014.9	4055.5	-9307.9	-9488.9
	(12200.3)	(12285.3)	(8845.1)	(8965.1)
Observations	2911	2873	2873	2873
Weights	Unweighted	Unweighted	Foster-Milusheva	Backbone weight
Sample	men	Weights nonmissing	Weights nonmissing	Weights nonmissing
Controls	all	all	all	drop 74 bari/fam size

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.26: Robustness to Weighting: Labor Productivity

	(1)	(2)	(3)	(4)
	labor_productivity	labor_productivity	labor_productivity	labor_productivity
Treat	61.84	390.5	-890.6	-979.4
	(1741.2)	(1744.8)	(2290.4)	(2285.9)
Treat × Born 1970-1977	1719.5	859.7	139.3	754.7
	(4305.4)	(4330.4)	(4127.0)	(4078.6)
Treat × Born 1977-1982	-10106.4**	-10733.4**	-6981.0	-6836.5
	(4567.1)	(4591.3)	(4599.5)	(4697.2)
Treat × Born 1982-1988	21.96	-215.0	-2206.4	-2245.5
	(3705.2)	(3738.4)	(4715.0)	(4739.4)
Observations	2911	2873	2873	2873
Weights	Unweighted	Unweighted	Foster-Milusheva	Backbone weight
Sample	men	Weights nonmissing	Weights nonmissing	Weights nonmissing
Controls	all	all	all	drop 74 bari/fam size

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table 3.27: Agriculture Outcomes: Individual-level

	(1) Used high-yield seeds	(2) Crop profit/acre	(3) Used high-yield seeds	(4) Crop profit/acre
Treat	0.0476**	25.29	0.0560*	-6.792
	(0.0213)	(36.47)	(0.0303)	(47.08)
Born 1970-1977 x Treat			-0.0200	21.08
			(0.0327)	(54.86)
Born 1977-1982 x Treat			-0.0427	113.2**
			(0.0328)	(48.79)
Born 1982-1988 x Treat			0.0132	46.13
			(0.0242)	(44.58)
Observations	8281	8281	8281	8281
R^2	0.028	0.027	0.028	0.028
1974 baseline controls	Y	Y	Y	Y
Weighted	1/# HH memb.	1/# HH memb.	1/# HH memb.	1/# HH memb.

Notes: Standard errors are clustered by pre-program village. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.28: Agriculture Outcomes: Individual-level

	(1) Used high-yield seeds	(2) Crop profit/acre	(3) Used high-yield seeds	(4) Crop profit/acre
Treat	0.0470	43.85	0.0611	14.72
	(0.0285)	(47.18)	(0.0420)	(55.74)
Born 1970-1977 x Treat			0.0145	35.53
			(0.0465)	(64.43)
Born 1977-1982 x Treat			-0.0729	88.76
			(0.0549)	(67.96)
Born 1982-1988 x Treat			-0.0309	41.97
			(0.0380)	(75.36)
Observations	7185	7185	7185	7185
R^2	0.050	0.045	0.052	0.046
1974 baseline controls	Y	Y	Y	Y
Weighted	1/# HH memb.× FM weight	1/# HH memb.× FM weight	1/# HH memb.× FM weight	1/# HH memb.× FM weight

Notes: Standard errors are clustered by pre-program village. Sample weights from Foster and Milusheva (2017). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.29: Enterprise Outcomes: Household-Level (Head Treat)

	(1) Num. business owned	(2) Any business loans	(3) Keep dedicated bank acct. for business	(4) Keep detailed accts. for business
Treat	0.0867	0.0353	0.00924	-0.00371
	(0.105)	(0.0229)	(0.0110)	(0.0267)
Observations	2510	2510	2510	2510
Mean of dep. var.				
Baseline controls				

	(1) Num. workers employed	(2) Tot profits	(3) Tot revenue	(4) Avg revenue per worker
Treat	-0.0773	-13573.5	73.23	428.5
	(0.200)	(11318.8)	(5069.9)	(428.5)
Observations	2510	2510	2510	2510
Mean if born 1970-77				
Mean if born 1977-82				
Mean if born 1982-88				

Notes: The table presents estimates of equation **??** at the household-level. Standard errors are clustered by pre-program village. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.30: Enterprise Outcomes: Household-Level (Frac Treat)

	(1)	(2)	(3)	(4)
	Num. business owned	Any business loans	Keep dedicated bank acct. for business	Keep detailed accts. for business
% HH treated	0.0890	0.0422*	0.0107	-0.00322
	(0.106)	(0.0220)	(0.0119)	(0.0284)
Observations	2510	2510	2510	2510
Mean of dep. var.				
Baseline controls				

	(1)	(2)	(3)	(4)
	Num. workers employed	Tot profits	Tot revenue	Avg revenue per worker
% HH treated	-0.0969	-12458.8	646.2	546.0
	(0.201)	(12478.4)	(5648.3)	(486.0)
Observations	2510	2510	2510	2510
Mean if born 1970-77				
Mean if born 1977-82				
Mean if born 1982-88				

Notes: The table presents estimates of equation **??** at the household-level. Standard errors are clustered by pre-program village. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.31: Agriculture Outcomes: Comparing Treatment Variables

	(1)	(2)	(3)	(4)
	=1 if used high-yield seed	=1 if used high-yield seed	Agr profit/acre	Agr profit/acre
Treat	0.0901**		83.49**	
	(0.0350)		(36.63)	
% HH treated		0.0852**		75.90**
		(0.0370)		(36.85)
Observations	2510	2510	2510	2510
R^2	0.036	0.034	0.045	0.044
1974 baseline controls	Y	Y	Y	Y
Weighted	Y	Y	Y	Y

Notes: Standard errors are clustered by pre-program village. Sample weights from Foster and Milusheva (2017). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.32: Agriculture Outcomes: Adding Embankment

	MHSS2 (2012-2014)		
	(1)	(2)	(3)
	=1 if household farms	Agr profit per acre	=1 if household farms
Treat	0.0413	48.53	0.0840**
	(0.0323)	(36.23)	(0.0379)
Embanked side, unprotected	-0.103**	-22.68	0.0455
	(0.0492)	(35.72)	(0.0456)
Non-embanked side	-0.0312	97.51***	0.0791**
	(0.0334)	(32.47)	(0.0361)
Observations	2273	2273	2273
Adjusted R^2	0.092	0.045	0.036
Baseline controls	Y	Y	Y
% chg. rel. to mean	6.4	17.1	40.4

Notes: Standard errors clustered at the baseline village level. Weights from Foster and Milusheva (2017). *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3.33: Enterprise Outcomes: Adding Embankment

	(1) Num. business owned	(2) Have business Loans	(3) Keep dedicated bank acct. for business	(4) Keep detailed accts. for business
PANEL A				
Treat	0.00293	0.0359	-0.0187*	-0.0291
	(0.0545)	(0.0220)	(0.0102)	(0.0262)
Treat × Born 1970-1977	-0.105	-0.0684**	0.0319	0.0354
	(0.0843)	(0.0345)	(0.0207)	(0.0348)
Treat × Born 1977-1982	0.129*	-0.0436	0.00788	0.00361
	(0.0771)	(0.0360)	(0.0203)	(0.0458)
Treat × Born 1982-1988	0.0971	-0.0237	0.0257	0.0380
	(0.0677)	(0.0243)	(0.0178)	(0.0360)
Embanked side, unprotected	-0.00670	0.0306*	-0.00779	0.0104
	(0.0489)	(0.0159)	(0.0102)	(0.0248)
Non-embanked side	-0.0388	0.0247	-0.00334	0.00902
	(0.0355)	(0.0152)	(0.00806)	(0.0186)
Observations	2752	2752	2752	2752
Mean if born 1970-77	0.69	0.16	0.03	0.17
Mean if born 1977-82	0.44	0.10	0.04	0.16
Mean if born 1982-88	0.33	0.04	0.03	0.10

	(1) Num. workers employed	(2) Tot. profits (taka)	(3) Revenue	(4) Revenue per worker
PANEL B				
Treat	-0.179	-18471.9	9179.5	1146.6
	(0.171)	(12832.5)	(6142.4)	(1905.0)
Treat × Born 1970-1977	0.00945	42028.5	-2148.0	1206.6
	(0.244)	(36297.4)	(9627.8)	(3712.3)
Treat × Born 1977-1982	0.199	31860.9	-37574.2***	-10332.0**
	(0.273)	(39972.9)	(14314.5)	(4625.4)
Treat × Born 1982-1988	0.356	55770.6*	-4348.9	-2005.7
	(0.226)	(28934.3)	(11578.2)	(3612.1)
Embanked side, unprotected	-0.00104	5928.1	4713.0	2679.1
	(0.220)	(21439.9)	(9064.0)	(3603.1)
Non-embanked side	-0.176	3954.3	162.8	-657.6
	(0.108)	(11226.2)	(3984.1)	(1305.3)
Observations	2752	2752	2752	2752
Mean if born 1970-77	1.2	69036.1	17456.3	10466.0
Mean if born 1977-82	0.8	56948.6	37200.9	15115.8
Mean if born 1982-88	0.5	33017.1	14608.2	6846.5

Notes: The table presents estimates of equation 3.1 at the individual-level. Means by age group are for the comparison group. Standard errors are clustered by pre-program village. *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Literature Cited

Daron Acemoglu and Simon Johnson. Disease and development: the effect of life expectancy on economic growth. *Journal of Political Economy*, 115(6):925–985, 2007.

David Atkin, Azam Chaudhry, Shamyla Chaudry, Amit K. Khandelwal, and Eric Verhoogen. Organizational barriers to technology adoption: Evidence from soccer-ball producers in Pakistan. *Quarterly Journal of Economics*, 132(3):1101–1164, 2017.

Tania Barham. Enhancing cognitive functioning: Medium-term effects of a health and family planning program in Matlab. *American Economic Journal: Applied Economics*, 4(1):245–73, 2012.

Tania Barham and Randall Kuhn. Staying for benefits the effect of a health and family planning program on out-migration patterns in Bangladesh. *Journal of Human Resources*, 49(4):982–1013, 2014.

Tania Barham, Randall Kuhn, and Patrick Turner. No place like home: Long-run impacts of early child health and family planning on labor and migration outcomes. Technical report, October 2019.

Tania Barham, Gisella Kagy, Brachel Champion, and Jena Hamadani. Early childhood health and family planning: Long-term and intergenerational effects on human capital. Technical report, 2021.

Hoyt Bleakley. Malaria eradication in the Americas: A retrospective analysis of childhood exposure. *American Economic Journal: Applied Economics*, 2(2):1–45, 2010.

Hoyt Bleakley and Fabian Lange. Chronic disease burden and the interaction of education, fertility, and growth. *The Review of Economics and Statistics*, 91(1):52–65, 2009.

Günther Fink and Felix Masiye. Health and agricultural productivity: Evidence from Zambia. *Journal of health economics*, 42:151–164, 2015.

Marshal F. Folstein, Susan E. Folstein, and Paul R. McHugh. "mini-mental state": a practical method for grading the cognitive state of patients for the clinician. *Journal of Psychiatric Research*, 12(3):189–198, 1975.

Andrew Foster and Sveta Milusheva. Household recombination, retrospective evaluation, and educational mobility over 40 years. 2017.

Andrew D. Foster and Mark R. Rosenzweig. Technical change and human-capital returns and investments: evidence from the green revolution. *The American Economic Review*, pages 931–953, 1996.

James J. Heckman. Skill formation and the economics of investing in disadvantaged children. *Science*, 312(5782):1900–1902, 2006.

Shareen Joshi and T. Paul Schultz. Family planning and women's and children's health: Long-term consequences of an outreach program in Matlab, Bangladesh. *Demography*, 50(1):149–180, 2013.

Michael A. Koenig, Mehrab Ali Khan, Bogdan Wojtyniak, John D. Clemens, Jyotsnamoy Chakraborty, Vincent Fauveau, James F. Phillips, Jalaluddin Akbar, and Uday S. Barua. Impact of measles vaccination on childhood mortality in rural Bangladesh. *Bulletin of the World Health Organization*, 68(4):441, 1990.

Jane Menken and James F. Phillips. Population change in a rural area of Bangladesh, 1967-87. *The Annals of the American Academy of Political and Social Science*, 510(1): 87–101, 1990.

James F. Phillips, Wayne S. Stinson, Shushum Bhatia, Makhlisur Rahman, and Jyotsnamoy Chakraborty. The demographic impact of the family planning–health services project in Matlab, Bangladesh. *Studies in Family Planning*, pages 131–140, 1982.

Mark Pitt and Mark R. Rosenzweig. Agricultural prices, food consumption, and the health and productivity of Indonesian farmers. *Agricultural household models: Extensions, applications and policy*, pages 153–82, 1986.

Gary Solon, Steven J. Haider, and Jeffrey M. Wooldridge. What are we weighting for? *Journal of Human Resources*, 50(2):301–316, 2015.

Willen van Schendel and Mahbubar Rahman. Gender and the inheritance of land: Living law in Bangladesh. *The Village in Asia Revisited*, pages 237–276, 1997.

David N. Weil. Accounting for the effect of health on economic growth. *Quarterly Journal of Economics*, 122(3):1265–1306, 2007.

World Bank. *World development report 2008: Agriculture for development*. The World Bank, 2007.