UC Berkeley
UC Berkeley Electronic Theses and Dissertations

Title
Essays in Development and Trade

Permalink
https://escholarship.org/uc/item/9hf0p725

Author
Galle, Simon

Publication Date
2016

Peer reviewed|Thesis/dissertation

eScholarship.org Powered by the California Digital Library
University of California

Essays in Development and Trade

by

Simon Galle

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Andrés Rodríguez-Clare, Co-Chair
Edward Miguel, Co-Chair
Benjamin Faber
Yuriy Gorodnichenko
Jeremy Magruder

Summer 2016

Essays in Development and Trade

Copyright 2016
by
Simon Galle

Abstract

Essays in Development and Trade

by

Simon Galle

Doctor of Philosophy in Economics

University of California, Berkeley

Andrés Rodríguez-Clare, Co-Chair

Edward Miguel, Co-Chair

This dissertation in development and trade explores the economic impact of liberalization and globalization. In the past few decades, as emerging economies such as India and China have opened up to world trade and liberalized their economies, these countries have experienced a surprisingly fast increase in GDP per capita. This is testament to the large benefits that can be reaped from globalization and liberalization. Paradoxically however, while globalization and liberalization should be celebrating their success stories, they are met by ever fiercer criticism, as is clear from rising opposition against free trade and an increasing resentment against globalization on both sides of the Atlantic, of which the recent Brexit referendum is but one example. These developments call for a more nuanced understanding of the benefits, but also of the downsides of globalization and liberalization, and this dissertation attempts to contribute to this understanding.

The first chapter of this dissertation develops a novel general-equilibrium model of the relationship between competition, financial constraints and misallocation. In the model, steady-state misallocation consists of both variable markups and capital wedges. The variable markups arise from Cournot-type competition, whereas the capital wedges result from the interaction of firm-level productivity volatility with financial constraints. Firms experience random shocks to their productivity and in response to positive productivity shocks they optimally grow their capital stock, subject to financial constraints. Competition plays a dual role in affecting misallocation. On the one hand, both markup levels and markup dispersion tend to fall with competition, which unambiguously improves allocative efficiency in a setting without financial constraints. On the other hand, in a setting with financial constraints, a reduction in markups is associated with slower capital accumulation, as the rate of self-financed investment shrinks. Thus, the positive impact of competition on steady-state misallocation is reduced by the presence of financial constraints.

The second chapter then tests the implications of the theoretical model from the first chapter using Indian plant-level panel data. The prediction that the firm-level speed of capital convergence falls with competition is confirmed for the full panel of manufacturing

plants in India's Annual Survey of Industries. This effect is particularly pronounced in sectors with higher levels of financial dependence. I also exploit natural variation in the level of competition, arising from the pro-competitive impact of India's 1997 dereservation reform on incumbent plants, and again confirm the qualitative predictions of the model.

The third chapter, which is joint work with Andrés Rodríguez-Clare and Moises Yi, develops and applies a framework to analyze the effect of trade on aggregate welfare as well as the distribution of this aggregate effect across different groups of workers. The framework combines a multi-sector gravity model of trade with a Roy-type model of the allocation of workers across sectors. The model predicts unequal distribution of the gains from trade as labor demand increases (decreases) for groups of workers specialized in export-oriented (import-oriented) sectors. The model generalizes the specific-factors intuition to a setting with labor reallocation, while maintaining analytical tractability for any number of groups and countries. We bring the model to the data using China's growth as a trade shock, where we define groups as German regions. First, we show that the model's structure accurately captures the empirical changes in regional income due to the China shock. Second, we structurally estimate the model's parameter that governs the distributional effects of the model. Counterfactual simulations show that this parameter implies sizable distributional implications of trade, with several groups losing from free trade. Finally, we measure the "inequality-adjusted" welfare effect of trade, which captures the full cross-group distribution of welfare changes in one measure. We find that inequality-adjusted gains from trade are larger than the aggregate gains for both countries, as between-group inequality falls with trade relative to autarky. Importantly, the opposite happens for the China shock.

That was very close to the truth,
but I don't think it's gonna make much sense.

David Foster Wallace

Escribo para el pueblo, aunque no pueda
leer mi poesía con sus ojos rurales.
Vendrá el instante en que una línea, el aire
que removió mi vida, llegará a sus orejas,
y entonces el labriego levantará los ojos,
el minero sonreirá rompiendo piedras,
el palanquero se limpiará la frente,
el pescador verá mejor el brillo
de un pez que palpitando le quemar las manos,
el mecánico, limpio, recién lavado, lleno
de aroma de jabón mirará mis poemas,
y ellos dirán tal vez: "Fue un camarada."

Pablo Neruda, Canto General

To my grandfathers, peter Wies and peter Albert.

Contents

Contents **ii**

List of Figures **iv**

List of Tables **v**

1 Competition, Financial Constraints and Misallocation: a Theoretical Analysis **1**
1.1 Introduction 1
1.2 Theory 3
1.3 Proof for Proposition 1 14
1.4 Concluding Remarks 23

2 Competition and Capital Convergence: Plant-level Evidence from Indian Manufacturing **24**
2.1 Introduction 24
2.2 Data 26
2.3 Stylized facts 27
2.4 Competition and capital convergence 31
2.5 Competition policy reform: dereservation 39
2.6 Concluding Remarks 45

3 Slicing the Pie: Quantifying the Aggregate and Distributional Effects of Trade **46**

3.1 Introduction 46
3.2 Theory 50
3.3 Data 64
3.4 Empirics 66
3.5 Counterfactual Simulations 72
3.6 Conclusion 81

Bibliography **82**

A Competition, Financial Constraints and Misallocation **89**
A.1 Labor market equilibrium 89
A.2 Model with young firms 91
A.3 Markup Measurement 95
A.4 Further stylized Facts 95
A.5 Event Study on Dereservation: Urban/Rural Distinction 98
A.6 Capital-labor ratio convergence 100

B Slicing the Pie **102**
B.1 Trade and Aggregate Sectoral Reallocation in Germany 102
B.2 US version 104
B.3 First-stages of estimation procedures 110
B.4 Extensions with home production 111

List of Figures

2.1 Productivity volatility and MRPK dispersion 29
2.2 Dereservation Event-study on Markups and Capital Growth 41

3.1 Germany moves to autarky: distribution of gains by region 73
3.2 Distribution of Gains by Region 74
3.3 Inequality-adjusted Gains from Trade 75
3.4 Relation between Import Competition and Earnings per Worker 76
3.5 Distribution of Gains by Region - $\forall s : \hat{T}_{China,s}^{1/\theta} = 5$ 77
3.6 Welfare-effects and changes in import-competition 79
3.7 Relation between $\ln \sum_s \pi_{igs} \hat{r}_{is}$ and κ 79
3.8 Inequality-Adjusted welfare-effects from the China shock 80
3.9 Relation between income per worker and $\ln \hat{I}_g$ - China shock 81

A.1 Dereservation Event-study on Markups and Capital Growth 99

B.1 Decomposition of Changes in Output Shares 103
B.2 Relation between Sectoral Output and Employment Shares 104
B.3 Decomposition of Changes in Output Shares - US 105
B.4 Relation between Sectoral Output and Employment Shares 106
B.5 Distributional Gains by Region - Autarky - US 107
B.6 Distributional Gains by Region - Autarky - US 107
B.7 Correlation between import-competition and earnings per worker 108
B.8 IGT - US 109
B.9 First stage for Table 3.4 110
B.10 First stage for Table 3.3 110

List of Tables

2.1 Capital growth as a Function of Age 30
2.2 MRPK Dispersion and Competition 31
2.3 Speed of MRPK convergence 37
2.4 Speed of Convergence for Young Plants 38
2.5 Speed of MRPK Convergence after Dereservation 43
2.6 Speed of Convergence for Young Plants after Dereservation 44

3.1 List of Industries 64
3.2 Labor Reallocation in Response to Trade Shock 68
3.3 Changes in import-competition and regional income per worker 69
3.4 Reallocation and regional income per worker 71
3.5 Germany's return to autarky: summary statics 72
3.6 Index of sectoral import competition 74
3.7 $\hat{W}_{ig}$ in Germany - $\forall s : \hat{T}_{China,s}^{1/\theta} = 5$ 77

A.1 MRPK dispersion: plant-level robustness 96
A.2 Difference in markups: urban versus rural 97
A.3 Capital-labor ratio: speed of convergence 101

B.1 Decomposition of Changes in Output Shares 104
B.2 Decomposition of Changes in Output Shares 106

Acknowledgments

The completion of this dissertation marks the end of my student life. While many view academic study mostly as a preparation for "real life", for me it *has been my life* for the past fourteen years. These fourteen years have shaped and often perturbed my outlook on the world and they have changed me as a person. I can then only hope I have indeed, in the spirit of the renaissance ideal, become "humanior." As I realize that I have continuously relied on others along my academic journey, I also become aware how I fail to accurately grasp the extent of my indebtedness. I cannot then expect to express my gratitude in a fully appropriate manner. Nevertheless, I will try.

First, I thank my advisor Andrés Rodríguez-Clare for his patient guidance and persistently insightful supervision during the past five years. Andrés has been a mentor in the full sense of the word and I could not have completed my PhD without him. I am also very grateful to Ted Miguel for his advice, ideas and support. His own work has inspired me to work in development economics, and in particular on questions at the core of economic development. Moreover, Ted has been central to my professional development, and working with him in Kenya has deeply influenced my outlook on development and growth.

Ben Faber and Yuriy Gorodnichenko deserve my gratitude for their invaluable advice and guidance. I admire Ben's research which combines a structural approach with innovative identification strategies. Needless to say then that his advice has been crucial for the completion of my dissertation. As for Yuriy's role as an advisor, "good wine needs no bush," and in his case the "wine" is extraordinary, both at an intellectual and at a personal level. I am also deeply grateful to Jeremy Magruder, Ben Handel and Cécile Gaubert for their suggestions, insightful comments and overall support, which has been central to my academic progress at different stages of the dissertation. Furthermore, I am indebted to seminar participants at BI Norwegian Business School, Bocconi, CREI, ECARES-ULB, ECORES Summer School, Edinburgh, Erasmus Rotterdam, HEC Lausanne, HEC Paris, Johns Hopkins SAIS, KULeuven, PEDL-CEPR, Surrey, Tilburg, UCLA Anderson and the Berkeley Development and Macro & Trade Seminars, the Trade and the Development Lunches and GEMS.

I started studying economics after completing my "licentiaat" in Philosophy. Economics didn't make sense at the time. The simplicity of the intellectual framework often times felt borderline repulsive and contrasted deeply with the rich diversity of perspectives in philosophy. It was Frans Spinnewyn who convinced me by his personal example about the relevance of economics. To this day, his authentic interest in the social dimension of economics continues to be an inspiration. Frans used to have verses from Neruda's Canto General on his office wall, and I cite these verses at the start of this dissertation to serve as a personal reminder as to what economics should be about.

During my master's at ULB, I had the privilege to have Bram De Rock as an advisor. Bram's guidance, encouragement and overall mentorship were absolutely crucial for my development as an economist. In addition to Bram's mentorship, the broader intellectual environment at ULB was conducive for me to apply to Berkeley. Mathias Dewatripont's course on asymmetric information at the start of the financial crisis, Gérard Roland's in-

troduction into institutional economics, and Philippe Weil's rigorous deconstruction of e.g. Ricardian Equivalence,... It all further convinced me that economics was the right choice for me. Both ULB and Berkeley have been an intellectual home, as my teachers at both institutions aim to incorporate a rich diversity of perspectives and challenge the mainstream views in economics. This is why I am grateful to have benefited from insightful teachers at Berkeley such as David Card, Brad DeLong, Barry Eichengreen, Fred Finan, Patrick Kline, Maury Obstfeld, Matthew Rabin, David Romer and Kenneth Train. I should also thank Vicky Lee and Patrick Allen for all the support they gave me and all the Berkeley graduate students. The department would be lost without Vicky and Patrick.

The past six years would not have been the same without my friends and office mates Juan-Pablo Atal, Pierre Bachas, and Dorian Carloni. Apparently, the worst office comes with the best office mates. It's been a true pleasure to work with my friend and coauthor Moises Yi on the "Slicing the Pie" paper, and I thank Moises in particular for being in charge of the relaxation activities these past years. It has also been a privilege to get to know Youssef Benzarti, Marco Schwarz, Alisa Tazhitdinova, and Torsten Walter. I will always treasure their friendship.

I feel fortunate for having been able to work with Kelly Zhang on our ethnicity paper. Especially my time in Nairobi would not have been the same without her. Then, I'm delighted I had the opportunity to meet with many inspiring people such as Marc Kaufmann, Yury Yatsynovych, David Silver, Attila Lindner, Hedvig Horvath, Yusuf Mercan, Asrat Tesfayesus, Miguel Almunia and Sandile Hlatshwayo. It was also a pleasure to be part of the trade research group, which has been crucial in the development of my dissertation.

At the homefront in Belgium, I feel very lucky to have such tremendous friends as Roel Faes, Tim Bauwens, Katrien Spruyt, Tine Vandendriessche and Matthias Somers. They were always present whenever I needed them, and I am sure we will remain friends for the rest of our lives. My family deserves a special thanks for always making me feel at home whenever I returned to Belgium, and I feel sorry for not having spent more time with them these past years. This holds in particular for my brother Michiel, Omayra, meter Judith, meter Simonne, tante Inge, my godchild Louis, Benot and Tom. Most importantly, my parents deserve enormous praise and infinite gratitude for always supporting me since I was born, but especially the past 6 years.

Finally, there are no words to express how grateful I feel to mijn liefje Fenella Carpena. Her loving care has always been present the past few years, and I could not have completed the PhD without her.

Chapter 1

Competition, Financial Constraints and Misallocation: a Theoretical Analysis

1.1 Introduction

Misallocation of resources has recently become a prominent explanation for cross-country differences in economic development. For instance, Hsieh and Klenow (2009) argue that misallocation, arising from the misalignment of marginal products across plants, could account for 40 to 60% of the difference in aggregate output per capita between the United States and India. This finding has sparked a debate on the main driving forces of the pattern in measured misallocation across countries. For instance, dispersion in the marginal revenue products of capital (MRPK), central to this paper's analysis, can be explained by either technological constraints, market imperfections or policy distortions.[1] Knowledge on the relative importance of these different underlying mechanisms matters to understand the potential level of macroeconomic efficiency gains from specific policy interventions.

This paper contributes to the above debate by investigating the relationship between competition, financial constraints and misallocation. Theoretically, existing work (Epifani

[1]Roughly speaking, capital misallocation is a function of the dispersion in marginal revenue products of capital (MPRK). As such it is a salient component of aggregate misallocation. Asker, Collard-Wexler, and De Loecker (2014) propose a model where such dispersion in MRPK is explained by adjustment costs in capital, which is a form of technological constraints. In this case, the dispersion in MRPK is the consequence of first-best optimization, and does not constitute a *mis*allocation of capital. In other settings measured dispersion in MRPK arises from market imperfections or policy distortions. In Midrigan and Xu (2014) and Moll (2014) the explanation for capital misallocation relies on market imperfections as firms' collateral constraints arise from imperfect financial markets. Restuccia and Rogerson (2008), in their seminal contribution to the misallocation literature, model misallocation as the result of firm-level variation in taxes or subsidies, with a non-competitive banking sector varying its interest rates for noneconomic reasons as a leading example. Restuccia and Rogerson (2013) provide a broader survey of the misallocation literature, while Buera, Kaboski, and Shin (2015) survey the literature on the macro-economic impact of financial constraints.

and Gancia, 2011; Peters, 2013) explains how in a setting with variable markups, competition reduces misallocation by decreasing dispersion in markups. While this channel is still present in my analysis, I demonstrate that financial constraints introduce a second, negative impact of competition on misallocation. Specifically, I show that competition slows down the capital growth rate of financially constrained firms and thereby capital wedges, resulting from the difference between the firm's optimal and actual capital levels, are amplified by competition. Intuitively, firm-level markups fall with the degree of competition, which lowers the scope for internally financed capital accumulation.

In the model, capital misallocation arises due to the interaction of productivity volatility and financial constraints. Productivity volatility implies that firms experience random shocks to their idiosyncratic levels of productivity. After a positive productivity shock, a firm will optimally choose to grow its capital stock, but the financial constraint will limit its ability to do so. Therefore, the scope for internal financing, governed by the level of markups in the industry, codetermines the speed of convergence to the firm's optimal capital level. Since competition reduces the level of markups, it will negatively affect a firm's speed of capital convergence in response to a positive productivity shock. This way, capital wedges are amplified by competition.

A related channel through which competition can negatively affect capital misallocation applies to young plants in particular. In an extension of the model, newborn firms are assumed to be undercapitalized and therefore financially constrained. As such, these firms will also rely on internal financing while converging to their optimal level of capital. This implies that competition again reduces the speed of capital convergence by reducing the scope for internal finance, thereby amplifying capital wedges.[2]

The theory builds on Midrigan and Xu (2014), who examine comparative statics for steady-state capital misallocation in a setting of imperfect competition. Since they employ simulation-based methods, the current paper theoretically contributes to the literature by providing an analytical solution for capital misallocation as a function of competition. The main focus of Midrigan and Xu (2014) is in quantifying the relative importance of barriers to entry versus collateral constraints for incumbent firms in shaping misallocation. This paper's focus on the comparative statics for competition is therefore complementary to their analysis.

Moll (2014) and Itskhoki and Moll (2015) also analyze capital misallocation analytically. However, they do so in a setting of perfect competition, whereas I study the impact of varying levels of imperfect competition. With its focus on the impact of competition on misallocation, this paper shares the orientation on policy with Itskhoki and Moll (2015), who study the impact of taxation policy on capital misallocation.[3] However, they focus

[2] As a preliminary empirical check, I provide evidence that the two fundamental sources underlying capital misallocation in the model - productivity volatility and the rate of arrival of newborn firms - are empirically salient in the Indian manufacturing sector.

[3] Generally in models with financial frictions, the first-best policy consists in removing such financial frictions. The focus on taxation policy in Itskhoki and Moll (2015) arises from a second-best perspective. The analysis of competition policy can also be understood from a second-best policy perspective.

on the role for policy along the transition path from an undercapitalized economy to the steady-state, whereas I analyze steady-state misallocation.

By examining the potential downsides of intensified competition, this paper complements papers that emphasize the beneficial impacts of competition on misallocation.[4] For instance, Peters (2013) argues that increased competition diminishes misallocation, as it reduces the dispersion in the distribution of markups. A second, well-established, beneficial impact of competition consists in reallocating labor from low productivity to high productivity firms. Here, Melitz (2003) studies the role of trade liberalization in improving the allocative efficiency of labor, and Akcigit, Alp, and Peters (2014) analyze constraints to such reallocation through competition in a Schumpeterian growth model with firm-level limits to delegation.

1.2 Theory

Setup of the economy

Agents The economy has two types of agents: workers and firm owners. The measure L of workers supplies labor inelastically, and each worker is hired at a wage w_t, where t indicates the time period. A worker's consumption c_{lt} is hand-to-mouth.

There is an exogenous, finite set M of firm-owners.[5] Firm-owner i has the following intertemporal preferences at time s:

$$U_{it} = \sum_{t=s}^{\infty} \beta^{t-s} d_{it}$$

Where β is the discount factor and d_{it} is firm-owner consumption.[6]

Production of varieties Each firm produces a variety i with a Cobb-Douglas production function, using capital k_{it} and labor l_{it} as inputs:

$$y_{it} = a_{it} k_{it}^{\alpha} l_{it}^{1-\alpha} \tag{1.1}$$

Productivity a_{it} follows a stochastic process over the state-space $a_{it} \in \{a_L, a_H\}$, where $a_L < a_H$.[7] Firm-level productivity volatility, arising from this stochastic path of a_{it}, will be central

[4] In the innovation literature, it is well-established that increasing competition can have both positive and negative impacts on aggregate output (see e.g. Aghion et al. (2005), Aghion, Akcigit, and Howitt (2013), and Gilbert (2006)). Also in the empirical micro-development literature there is work that studies the downsides of competition (Macchiavello and Morjaria, 2015). However, in the misallocation literature, the downsides of misallocation have been understudied.

[5] The main comparative statics within the model will be on M, as modifying the degree of competition. Having M as exogenous simplifies the analytical solution of the model. In a simulation-based methodology, as employed by Midrigan and Xu (2014), one can endogenize the degree of competition.

[6] The simplifying assumption of linear firm-owner preferences will prove useful in the analytical derivation of a global solution for the firm-level path of capital.

[7] I follow Midrigan and Xu (2014) by assuming that in period t, the firm is informed about the distribution of a_{it+1}.

in the analysis of steady-state firm-dynamics in section 1.2. Importantly, capital is a dynamic input, subject to the equation of motion:

$$k_{it+1} = x_{it} + (1-\delta)k_{it}$$

with investment x_{it} taking place, and being financed at the end of period t. The decision about labor l_{it+1} is also made in period t, i.e. at the same time the decision on k_{it+1} is made, but labor l_{it+1} is only paid at the end of period $t+1$.[8]

Demand Investment x_{it}, workers' consumption c_{lt} and firm-owner consumption d_{it} all consist of shares of the final good Q_t, which is composed of varieties q_{it}:

$$Q_t \equiv M^{1-\frac{1}{\eta}} \left[\sum_{i=1}^{M} q_{it}^{\eta} \right]^{\frac{1}{\eta}} \tag{1.2}$$

where $M^{1-\frac{1}{\eta}}$ eliminates taste-for-variety (Blanchard and Kiyotaki, 1987).[9] [10] This expression for the composite good implies that firms face the following demand function q_{it}:

$$q_{it} = \left(\frac{p_{it}}{P_t} \right)^{-\frac{1}{1-\eta}} M^{-\frac{1}{\eta}} \left[\sum_{i=1}^{M} q_{it}^{\eta} \right]^{\frac{1}{\eta}} \tag{1.3}$$

where p_{it} is the price of variety i and P_t is the price of the final good:

$$P_t^{-\frac{\eta}{1-\eta}} \equiv \frac{1}{M} \sum_{i=1}^{M} p_{it}^{-\frac{\eta}{1-\eta}} \tag{1.4}$$

Financial constraint The above implies that firms face the following period-by-period budget constraint, where z_{it} is wealth at the end of period t: $z_{it} \equiv p_{it}y_{it} - w_t l_{it} + P_t(1-\delta)k_{it}$.

$$P_t(k_{it+1} + d_{it}) \leq z_{it} \tag{1.5}$$

The financial constraint implies that consumption d_{it} cannot be negative:

$$d_{it} \geq 0 \tag{1.6}$$

[8] The assumption of labor and capital being decided simultaneously, will simplify the optimization problem.

[9] This expression for the final good is employed by Jaimovich (2007) in a setting with variable markups, and it allows to restrict attention to the competitive effects of varying M, and ignore the taste-for-variety effects. Bénassy (1996) generalizes the idea of de-linking consumption-side taste-for-variety and firm-level market power.

[10] There is one sector, and Q_t is the composite good of that sector. Note that it should be straightforward to extend this to a multi-sector case when preferences are Cobb-Douglas across sectors, as expenditure shares are constant across sectors in that case.

Firm's problem

Market structure and firm problem I follow Atkeson and Burstein (2008) by assuming that each period, firms play a one-period game of quantity competition.[11] Specifically, each firm i sets a quantity y_{it+1} for sale, conditional on the quantities chosen by the other firms in the economy. As discussed in the previous subsection, firms make decisions about l_{it+1}, k_{it+1} in period t, knowing a_{it+1} and given the budget constraint $P_t(k_{it+1} + d_{it}) \leq z_{it}$. Therefore, any firm i's optimal decisions are $k_{it+1}\left(a_{it+1}, z_{it}, \mathbf{y_{-it+1}}\right)$, $l_{it+1}\left(a_{it+1}, z_{it}, \mathbf{y_{-it+1}}\right)$, where (a_{it+1}, z_{it}) characterizes the state for firm i and $\mathbf{y_{-it+1}}$ is the vector of decisions on y_{jt+1} for all $j \neq i$. Through the production function (1.1), the choice of k_{it+1}, l_{it+1} determines y_{it+1} and thereby $p_{it+1}(y_{it+1}, \mathbf{y_{-it+1}})$ as firms incorporate the demand function (1.3) into their optimization. As such, this setting entails the following intertemporal problem for the firm, where $\pi_{it}(k_{it}, l_{it}, \mathbf{y_{-it}}) \equiv p_{it}(y_{it}, \mathbf{y_{-it}})y_{it} - w_t l_{it}$:

$$\begin{aligned}\max_{d_{it}, k_{it+1}, l_{it+1}} \mathcal{L} = &\sum_{t=s}^{\infty} E_s\left[\beta^{t-s} d_{it}\right] + \\ &\sum_{t=s}^{\infty} E_s\left[\lambda_{it}\left(\pi_{it}(k_{it}, l_{it}, \mathbf{y_{-it}}) + P_t\left[(1-\delta)k_{it} - k_{it+1} - d_{it}\right]\right) + \Phi_{it}(d_{it})\right]\end{aligned} \tag{1.7}$$

Since each firm's decision on y_{it+1} depends on $(a_{it+1}, z_{it}, \mathbf{y_{-it+1}})$, $\mathbf{y_{it+1}}$ will be determined by $F(a(t+1), z(t))$, the joint distribution of a_{it+1} and z_{it}, and by the conditions in the labor and goods market implied by M, L.

$$\begin{aligned}k_{it+1}\left(a_{it+1}, z_{it}, F(a(t+1), z(t)), M, L\right) \\ l_{it+1}\left(a_{it+1}, z_{it}, F(a(t+1), z(t)), M, L\right)\end{aligned} \tag{1.8}$$

The optimal choices in (1.8) determine $p_{it+1}\left(a_{it+1}, z_{it}, F(a(t+1), z(t)), M, L\right)$, and given the firm's marginal cost thereby also determine the markup μ_{it+1}

$$\mu_{it+1}\left(a_{it+1}, z_{it}, F(a(t+1), z(t)), M, L\right) = \frac{\varepsilon_{it+1}\left(a_{it+1}, z_{it}, F(a(t+1), z(t)), M, L\right) - 1}{\varepsilon_{it+1}\left(a_{it+1}, z_{it}, F(a(t+1), z(t)), M, L\right)} \tag{1.9}$$

where the demand elasticity ε_{it} is:

$$\varepsilon_{it+1}\left(a_{it+1}, z_{it}, F(a(t+1), z(t)), M, L\right) = -\frac{1}{1-\eta} + \left(\frac{\eta}{1-\eta}\right)\frac{y_{it+1}^{\eta}}{\sum_i y_{it+1}^{\eta}} \tag{1.10}$$

Labor optimization The first-order condition for labor is standard:

$$E_s\left[\frac{\partial \pi_{it}(k_{it}, l_{it}, F(a(t+1), z(t)), M, L)}{\partial l_{it}}\right] = 0 \tag{1.11}$$

[11] I will assume that strategic interaction of firms is only within-period.

Intertemporal optimization Now I derive the first-order conditions for the dynamic part of the problem. Start with the first-order condition for d_{it}.

$$\frac{\partial \mathcal{L}}{\partial d_{it}} = \beta^{t-s} + E_s[-\lambda_{it} P_t + \Phi_{it}] = 0$$

Which implies the following condition:

$$\beta^{t-s} + E_s[\Phi_{it}] = E_s[\lambda_{it} P_t] \tag{1.12}$$

Then, the first-order condition for k_{it+1} implies:

$$E_s\left[\lambda_{it} P_t\right] = E_s\left[\lambda_{it+1} P_{t+1}\left((1-\delta) + \frac{1}{P_{t+1}} \frac{\partial \pi_{it+1}(k_{it+1}, l_{it+1}, F(a(t+1), z(t)), M, L)}{\partial k_{it+1}}\right)\right] \tag{1.13}$$

Decision rules for capital and consumption

Capital and consumption The combination of (1.12) and (1.13) allows me to find the decision rules for d_{it}, k_{it+1}. Taking the perspective of period $s = t$, there are then two cases, either $\Phi_{it} > 0$ or $\Phi_{it} = 0$.

- **Case 1** When $\Phi_{it} = 0$, then k_{it+1} is optimally set such that:[12]

$$1 = E_t\left[\lambda_{it+1} P_{t+1}\left((1-\delta) + \frac{1}{P_{t+1}} \frac{\partial \pi_{it+1}(k_{it+1}, l_{it})}{\partial k_{it+1}}\right)\right] \tag{1.14}$$

  And consumption $d_{it} = \frac{\pi_{it}(k_{it}, l_{it})}{P_t} - x_{it}$.

- **Case 2** When $\Phi_{it} > 0$, then $d_{it} = 0$ and the path of capital is determined by the budget constraint: $k_{it+1} = \frac{\pi_{it}(k_{it}, l_{it})}{P_t} + (1-\delta) k_{it}$.

Output and markup The above decision rules also imply an output decision for both cases.

- **Case 1** When $\Phi_{it} = 0$, then firms in period t solve the following system of decision rules regarding period $t+1$:

$$E_t\left[\lambda_{it+1} \frac{\partial \pi_{it+1}(k_{it+1}, l_{it})}{\partial k_{it+1}}\right] = 1 - E_t\left[\lambda_{it+1} P_{t+1}(1-\delta)\right]$$
$$\frac{\partial \pi_{it+1}(k_{it+1}, l_{it+1})}{\partial l_{it+1}} = 0$$

[12] In case $E_t[\lambda_{t+1} P_{t+1}] = \beta$, i.e. when $E_t[\Phi_{t+1}] = 0$, then (1.14) simplifies to $\frac{\partial \pi_{it+1}(k_{it+1}, l_{it})}{\partial k_{it+1}} = P_{t+1}\left(\frac{1}{\beta} + \delta - 1\right)$.

- **Case 2**: When $\Phi_{it} > 0$, then the optimal labor choice l_{it+1} is chosen conditional on $k_{it+1} = \frac{\pi_{it}(k_{it},l_{it})}{P_t} + (1-\delta)k_{it}$.

Given the decision on k_{it+1}, l_{it+1}, the output y_{it+1} is determined due to the production function (1.1). Then, given (1.3), this determines the price p_{it+1} of the firm. This pricing decision simultaneously implies a decision on the markup in (1.9), given the firm's marginal cost.

Steady state equilibrium

An *equilibrium* consists of a set of prices P_t, w_t, p_{it}, a set of consumption $d_{it}(a_{it+1}, z_{it}, F(a(t+1), z(t)))$, capital $k_{it+1}(a_{it+1}, z_{it}, F(a(t+1), z(t)))$ and labor $l_{it}(a_{it}, z_{it-1}, F(a(t), z(t-1)))$ decisions by firm-owners and consumption by workers $\frac{w_t}{P_t}L$ that satisfy

- the labor market clearing condition

$$L = \sum_{i=1}^{M} l_{it} \tag{1.15}$$

- the goods market clearing condition

$$Q_t = \sum_{i=1}^{M}(x_{it} + d_{it}) + \int_{l \in L} c_{lt} dl \tag{1.16}$$

- the optimality conditions (1.11), (1.13) for each firm i, conditional on the choices of l_{jt}, k_{jt} of all firms $j \neq i$.

- market-clearing for each variety i: $y_{it} = q_{it}$, satisfying (1.3)

- the equalized budget constraint $P_t(k_{it+1} + d_{it}) = z_{it}$, and the financial constraint $d_{it} \geq 0$.

To solve this equilibrium, I can pick as numeraire $w_t = 1$, and P_t is a function of the individual prices as in (1.4). Next, y_{it} is determined by k_{it}, l_{it}, a_{it}, where a_{it} is exogenous. Satisfying (1.3) implies that p_{it} is given by choice of y_{it}. Finally, l_{it}, k_{it}, d_{it} are determined by (1.11), (1.13) and the budget constraint (1.5), as explained in section 1.2. Since there are M firms, this then is a system of Mx3 equations with Mx3 unknowns.

A *steady state equilibrium* is an equilibrium that satisfies for all t[13] :

[13] Moll (2014) employs a similar definition of a steady state equilibrium.

$$\begin{aligned} K_t &= K, \\ \frac{P_t}{w_t} &= \frac{P}{w}, \\ F(a(t+1), z(t)) &= F(a', z) \end{aligned} \tag{1.17}$$

A first implication of this definition of the steady state, is that $H(a(t), k(t)) = H(a, k)$, i.e. the joint distribution of productivities and capital will be stable.[14]The reason is that capital choice is determined by $F(a(t+1), z(t))$: $k_{it+1}(a_{it+1}, z_{it}, F(a(t+1), z(t)))$. A second implication is that aggregate output will be stable as well: $Q_t = Q$.

Analysis of the steady state

Section 1.2 implies that in steady state each firm's decisions depend on $F(a+1, z)$. Here, the wealth distribution is endogenous, whereas the distribution of productivities is exogenously determined. Since the distribution of wealth is a function of $H(a, k)$, I focus on examining this joint distribution of productivities and capital in steady state. To this end, I will start by characterizing the firm's decision rules for capital and labor in steady state.

Labor and capital decisions in steady state

It will be convenient to characterize the solution to the firm's optimization problem by taking the perspective of the cost-minimization problem given the optimal markup characterized in (1.9).[15] As such, the cost-minimization problem implies the following optimal labor demand in steady state:

$$l_{it} = \left(\frac{(1-\alpha)}{\mu_{it}} \frac{P}{w} \left(\frac{Q}{M} \right)^{1-\eta} a_{it}^{\eta} k_{it}^{\alpha\eta} \right)^{\frac{1}{1+\alpha\eta-\eta}} \tag{1.18}$$

For the capital choice, as is clear from section 1.2, there are two cases: either $\Phi_{it} = 0$, or $\Phi_{it} > 0$.

Unconstrained firms First consider the case where a firm has $\Phi_{it} = 0$. In that case, the optimality condition in (1.13), together with (1.18) implies that

$$k_{it}^{*} = \mu_{it}^{\frac{1}{\eta-1}} a_{it}^{\frac{\eta}{1-\eta}} \frac{Q}{M} \left(\frac{P(1-\alpha)}{w} \right)^{\frac{\eta-\alpha\eta}{1-\eta}} \left(\frac{\alpha}{r_{it}} \right)^{\frac{1+\alpha\eta-\eta}{1-\eta}} \tag{1.19}$$

where $r_{it} \equiv \left(\frac{1}{\beta} + \delta - 1 \right) - \xi_{it}$.[16]

[14]I am assuming here that the productivity volatility process is such that it allows for a stable $H(a, k)$.

[15]Jaimovich (2007) also employs the cost-minimization approach to characterize the solution to the firm problem, and as such, the optimality conditions are closely related to the ones found in that paper.

[16]When $E_t[\Phi_{t+1}] = 0$, then $\xi_{it} = 0$, otherwise $\xi_{it} > 0$.

Constrained firms When the financial constraint binds, i.e. $\Phi_{it} > 0$. Capital grows according to the budget constraint. Specifically, I show in section 1.3 that:

$$k_{it+1} = (1-\delta)k_{it} + [\mu_{it} - (1-\alpha)] \frac{w_t}{P_t(1-\alpha)} l_{it} \tag{1.20}$$

Distribution and dynamics for firm-level capital

Given the expressions for k^*_{it}, and the path for capital of constrained firms in (1.20), I now characterize $H(a,k)$. First, consider the firms with $a_{it} = a_L$. In steady state, these firms cannot have $\Phi_{it} > 0$[17], and therefore these firms have $k_{it} = k^*_L$, the optimal level of k_{it} for low productivity firms. Note that $k_{it}(a_L) > k^*_L$ violates the firm's optimality conditions, as firms consume any capital in excess of k^*_L, and thereby satisfy the decision rule for capital in equation (1.19).

Second, there are the firms with $a_{it} = a_H$. For these firms, either $\Phi_{it} = 0$, or $\Phi_{it} > 0$. When $\Phi_{it} = 0$, then these firms have $k_{it} = k^*_H$. When $\Phi_{it} > 0$, then $k_{it} = G_\tau k^*_L$, where $\tau = t - s$,

$$G_\tau \equiv \Pi_{r=s}^{s+\tau}(1+g_r) \tag{1.21}$$

and

$$g_r \equiv \frac{k_{r+1}}{k_r} - 1;\ s \equiv \max r \text{ s.t. } a_{ir+1} = a_H \& a_{ir} = a_L$$

Here, k_{r+1} is determined by (1.20), for any firm i with capital level k_r. In words, k_{it} is determined by the cumulative capital growth G_τ since the firm's most recent positive productivity shock.

Capital of unconstrained firms Following (1.19), the optimal values for capital k^*_L, k^*_H are:

$$\begin{aligned} k^*_L &= \left(\frac{a_L^\eta}{\mu_L}\right)^{\frac{1}{1-\eta}} \left(\frac{\alpha}{r_L}\right)^{\frac{1+\alpha\eta-\eta}{1-\eta}} \left(\frac{P(1-\alpha)}{w}\right)^{\frac{\eta-\alpha\eta}{1-\eta}} \frac{Q}{M} \\ k^*_H &= \left(\frac{a_H^\eta}{\mu_H}\right)^{\frac{1}{1-\eta}} \left(\frac{\alpha}{r_H}\right)^{\frac{1+\alpha\eta-\eta}{1-\eta}} \left(\frac{P(1-\alpha)}{w}\right)^{\frac{\eta-\alpha\eta}{1-\eta}} \frac{Q}{M} \end{aligned} \tag{1.22}$$

Where μ_L, μ_H, characterized further in section 1.2, are the optimal level of markups for the respective firms. Furthermore, $r_H = \frac{1}{\beta} + \delta - 1$ since $E_t[\Phi_{it}] = 0$ for all firms with $a_{it} = a_H$ and $\Phi_{it} = 0$. Next, r_L is the value for r_{it} for all firms with $a_{it} = a_L$. Since for firms with $a_{it} = a_H$, the level of capital depends on G_τ, the value of Φ_{it} is also determined by τ, i.e. the number of periods since the most recent productivity shock. The above entails that the following lemma holds.

[17]Suppose this is not the case and there is at least one firm with $a_{it} = a_L \,\&\Phi_{it} > 0$. Then for all firms i with $a_{it} = a_L \,\&\Phi_{it} > 0$, $k_{it+1}(a_{it+1}, z_{it}, F(a+1,z)) > k_{it}$. Since these firms' This then violates the property of the steady state that $F(a(t+1), k(t)) = F(a', z)$.

Lemma 1. *Steady state $H(a,k)$ is determined by:*

- $a_i = a_L$, *then* $k_i = k_L^*$
- $a_i = a_H$ *then* $\forall i$ *with* $\tau = t - s$, *where* $s = \max r$ *s.t.* $a_{ir+1} = a_H \& a_{ir} = a_L$*:*
  - *if* $\Phi_\tau = 0$, *then* $k_{i\tau} = k_H^*$
  - *if* $\Phi_\tau > 0$, *then* $k_{i\tau} = G_\tau k_L^*$

Distribution of markups

Now, I characterize the distribution of markups. First, the markups for the unconstrained firms follow directly from (1.9), (1.10) and Lemma 1.

$$\begin{aligned}\mu_L(a_L, k_L^*, H(a,k), M) &\equiv \frac{1 - M^{\eta-1}\eta\frac{(a_L(k_L^*)^\alpha(l_L^*)^{1-\alpha})^\eta}{Q^\eta}}{\eta\left(1 - M^{\eta-1}\frac{(a_L(k_L^*)^\alpha(l_L^*)^{1-\alpha})^\eta}{Q^\eta}\right)} \\ \mu_H(a_H, k_H^*, H(a,k), M) &\equiv \frac{1 - M^{\eta-1}\eta\frac{(a_H(k_H^*)^\alpha(l_H^*)^{1-\alpha})^\eta}{Q^\eta}}{\eta\left(1 - M^{\eta-1}\frac{(a_H(k_H^*)^\alpha(l_H^*)^{1-\alpha})^\eta}{Q^\eta}\right)}\end{aligned} \tag{1.23}$$

Constrained firms For constrained firms, we know that $k_{it} = G_\tau k_L^*$, and the markup for these firms can be written as:

$$\mu_\tau(a_H, G_\tau k_L^*, H(a,k), M) \equiv \frac{1 - M^{\eta-1}\eta\frac{(y(a_H, G_\tau k_L^*), F(a,k))^\eta}{Q^\eta}}{\eta\left(1 - M^{\eta-1}\frac{(y(a_H, G_\tau k_L^*), F(a,k)))^\eta}{Q^\eta}\right)} \tag{1.24}$$

Together (1.23), (1.24), characterize the distribution of markups.

Capital wedges

Next, I analyze the capital wedges ω_{it}, which will be important in the analysis of aggregate TFP. The capital wedges are implicitly defined in the following way:

$$k_{it} = \left(\frac{a_{it}^\eta}{\mu_{it}}\right)^{\frac{1}{1-\eta}} \left(\frac{\alpha}{\omega_{it}}\right)^{\frac{1+\alpha\eta-\eta}{1-\eta}} \left(\frac{P(1-\alpha)}{w}\right)^{\frac{\eta-\alpha\eta}{1-\eta}} \frac{Q}{M} \tag{1.25}$$

where $\omega_{it} = r_L, r_H$ for unconstrained firms with productivities a_L, a_H respectively, and $\omega_{it} > r_H$ for constrained firms. For these constrained firms, I combine equations (1.22) and (1.25), to express the capital wedge for any period τ:[18]

[18]The expression is found after simplifying $\omega_{it} = \alpha(G_{it,s}k_L^*)^{-\frac{1-\eta}{1+\alpha\eta-\eta}}\left[\frac{a_{it}^\eta}{\mu_{it}}\left(\frac{Q_t}{M}\right)^{1-\eta}\left(\frac{P_t(1-\alpha)}{w_t}\right)^{\eta-\eta\alpha}\right]^{\frac{1}{1+\alpha\eta-\eta}}$

$$\omega_\tau = G_\tau^{-\frac{1-\eta}{1+\alpha\eta-\eta}} \left[\frac{a_H^\eta}{a_L^\eta}\frac{\mu_L}{\mu_\tau}\right]^{\frac{1}{1+\alpha\eta-\eta}} r_L \tag{1.26}$$

Note that: $\max_t \omega_\tau = \omega_1 = G_1^{-\frac{1-\eta}{1+\alpha\eta-\eta}} \left[\frac{a_H^\eta}{a_L^\eta}\frac{\mu_L}{\mu_1}\right]^{\frac{1}{1+\alpha\eta-\eta}} r_L$. Hence the distribution of ω_τ for firms with $a_{it} = a_H$, has a range $[r_H, \omega_1]$.

Lemma 2. *In steady state, the distribution of capital wedges is:*

- *For firms with $a_{it} = a_L$, $\omega_{it} = r_L$*
- *For firms with $a_{it} = a_H$:*
  - *When $\Phi_\tau = 0$, $\omega_{it} = r_H$*
  - *When $\Phi_\tau > 0$, $\omega_\tau(G_\tau, \mu_\tau) = G_\tau^{-\frac{1-\eta}{1+\alpha\eta-\eta}} \left[\frac{a_H^\eta}{a_L^\eta}\frac{\mu_L}{\mu_\tau}\right]^{\frac{1}{1+\alpha\eta-\eta}} r_L$*

Aggregates for output, capital and TFP

Aggregate output In appendix A.1, I show that

$$Q = TFPK^\alpha L^{1-\alpha} \tag{1.27}$$

where TFP is aggregate productivity and K is aggregate capital.

TFP I now characterize TFP. In appendix A.1, I derive equation (A.2), which is the explicit function for TFP. It is clear from that equation, that TFP is a function of the joint distribution of productivities, markups and capital wedges ω_{it}. Since the capital wedges are a function of a_{it}, k_{it}, I can use Lemma 1 and equations (1.23),(1.24), to characterize TFP as:

$$TFP = F_{TFP}(H(a,k), M) \tag{1.28}$$

Aggregate capital Given Lemma 1, aggregate capital $K_t = \sum_{i=1}^M k_{it}$ can in steady state be expressed as:

$$K = M\left[Prob(a_{it} = a_L)k_L^* + \sum_{\tau=1}^{\infty} Prob(a_{it} = a_H \& s = t-\tau)G_\tau k_L^*\right]$$

After substituting in the value for k_L^*, and using $Q = TFPK^\alpha L^{1-\alpha}$. We find: [19]

[19] Specifically:

$$K = Q\left(\frac{P(1-\alpha)}{w}\right)^{\frac{\eta-\alpha\eta}{1-\eta}} \alpha^{\frac{1+\alpha\eta-\eta}{1-\eta}} \left(\frac{a_L^\eta}{\mu_L r_L^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}} \left[Prob(a_{it} = a_L) + \sum_{\tau=1}^{\infty} Prob(a_{it} = a_H \& s = t-\tau)G_\tau\right]$$

$$K^{1-\alpha} = TFPL^{1-\alpha}\left(\frac{P(1-\alpha)}{w}\right)^{\frac{\eta-\alpha\eta}{1-\eta}}\alpha^{\frac{1+\alpha\eta-\eta}{1-\eta}}\left(\frac{a_L^{\eta}}{\mu_L r_L^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}} \left[Prob(a_{it}=a_L)+\sum_{\tau=1}^{\infty}Prob(a_{it}=a_H \& s=t-\tau)G_{\tau}\right] \tag{1.29}$$

Labor Market clearing

Since there are two markets, by Walras' Law, general equilibrium is realized when the labor market clears. Labor demand, given in equation (1.18), from all firms has to equal labor supply L:

$$L=\sum_{i=1}^{M}\left(\frac{(1-\alpha)}{\mu_{it}}\frac{P}{w}\left(\frac{Q}{M}\right)^{1-\eta}a_{it}^{\eta}k_{it}^{\alpha\eta}\right)^{\frac{1}{1+\alpha\eta-\eta}}$$

In appendix A.1, this equation is derived further. Then, notice that labor market clearing is realized for the following $\frac{P}{w}$:

$$\frac{P}{w}=\left(\frac{L}{K}\right)^{\alpha}\frac{\Omega^{\eta-\alpha\eta-1}}{(1-\alpha)\left(\frac{TFP}{M}\right)^{1-\eta}} \tag{1.30}$$

where

$$\Omega\equiv\left[\sum_{i=1}^{M}\left(\frac{a_{it}^{\eta}}{\mu_{it}}\left(\frac{\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}}}{\sum_{i=1}^{M}\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}}}\right)^{\alpha\eta}\right)^{\frac{1}{1+\alpha\eta-\eta}}\right] \tag{1.31}$$

Like TFP, Ω is a function of the joint distribution of productivities, markups and capital. In a context with monopolistic competition, i.e. without variable markups, this condition would not exist.

In short, the above implies that the labor market clearing equation can be written as:

$$\frac{P}{w}=F_L(M,L,K,TFP,\Omega) \tag{1.32}$$

Summary of steady-state equilibrium

The nature of the steady-state equilibrium will be determined by the following elements:

- $H(a,k)$, the joint distribution of a_{it}, k_{it}, characterized in Lemma 1
- The distribution of markups, characterized in equations (1.23), (1.24)

- Aggregate TFP, characterized in (1.28)
- Aggregate capital, characterized in (1.29)
- The factor-price ratio, determined in the labor-market-equilibrium condition in (1.32)
- Ω, characterized in (1.31).

In the comparative-statics exercise that now follows, I describe how the steady-state variables change with M. A crucial role there will be played by the comparative statics on G_τ, which is a crucial determinant of the distribution of capital.

Comparative statics on competition

In the theoretical appendix sections, I demonstrate the following proposition on the comparative statics for M:

Proposition 1. *For any $M' > M$, and for unconstrained firm-types L, H, and for constrained firms in period $\tau > 0$:*

- *Markup levels fall with M:*

$$\mu'_L < \mu_L \,;\, \mu'_H < \mu_H \,;\, \mu'_\tau < \mu_\tau$$

- *Markup dispersion falls with M:*

$$\frac{\mu'_H}{\mu'_L} < \frac{\mu_H}{\mu_L} \,;\, \frac{\mu'_\tau}{\mu'_L} \leq \frac{\mu_\tau}{\mu_L}$$

- *Capital wedges worsen with M:*

$$\omega'_\tau \geq \omega_\tau$$

and $(\Phi_\tau > 0) \implies (\omega'_\tau > \omega_\tau)$

The proposition demonstrates the dual role of competition in an environment with both variable markups and financial constraints. On the one hand, markup misallocation improves, since both markup levels and markup dispersion fall with M. On the other hand, misallocation due to capital wedges worsens due to competition. Since the latter effect is absent in a setting without financial constraints while the former is not, the welfare gains from competition tend to be lower in a setting with financial constraints compared to a setting without financial constraints.

1.3 Proof for Proposition 1

Proposition 1 has three components and in this section I provide the proof for each of the three components. It will be convenient to first focus on the third component, namely the relation between capital wedges and competition.

Overview of the proof

First, I demonstrate what the sufficient conditions are for the proposition's statement for capital wedges, namely that $\forall \tau : \frac{d\omega_\tau}{dM} \geq 0$ and $(\Phi_\tau > 0) \implies \frac{d\omega_\tau}{dM}$. To demonstrate this, I start from equation (1.26), which for convenience is reiterated here:

$$\omega_\tau = G_\tau^{-\frac{1-\eta}{1+\alpha\eta-\eta}} \left[\frac{a_H^\eta}{a_L^\eta} \frac{\mu_L}{\mu_\tau} \right]^{\frac{1}{1+\alpha\eta-\eta}} r_L$$

Therefore:

$$\begin{aligned}
\frac{d\omega_\tau}{dM} =& - \frac{1-\eta}{1+\alpha\eta-\eta} G_\tau^{-\frac{1-\eta}{1+\alpha\eta-\eta}-1} \frac{dG_\tau}{dM} \left[\frac{a_H^\eta}{a_L^\eta} \frac{\mu_L}{\mu_\tau} \right]^{\frac{1}{1+\alpha\eta-\eta}} r_L \\
&+ \frac{1}{1+\alpha\eta-\eta} \left(\frac{\mu_L}{\mu_\tau} \right)^{\frac{1}{1+\alpha\eta-\eta}-1} \frac{d(\frac{\mu_L}{\mu_\tau})}{dM} G_\tau^{-\frac{1-\eta}{1+\alpha\eta-\eta}} \left[\frac{a_H^\eta}{a_L^\eta} \right]^{\frac{1}{1+\alpha\eta-\eta}} r_L \\
&+ G_\tau^{-\frac{1-\eta}{1+\alpha\eta-\eta}} \left[\frac{a_H^\eta}{a_L^\eta} \frac{\mu_L}{\mu_\tau} \right]^{\frac{1}{1+\alpha\eta-\eta}} \frac{dr_L}{dM}
\end{aligned}$$

In this preliminary version of the paper, I assume that $\frac{dr_L}{dM} = 0$. A sufficient condition for $\frac{d\omega_\tau}{dM} \geq 0$ and $(\Phi_\tau > 0) \implies \frac{d\omega_\tau}{dM}$ to hold, is then the following two conditions hold:

- $\forall \tau > 0 : ((\frac{dG_\tau}{dM} \leq 0) \wedge (\Phi_\tau > 0)) \implies \frac{dG_\tau}{dM} < 0$
- $\forall \tau > 0 : \frac{d(\frac{\mu_L}{\mu_\tau})}{dM} > 0$

Here is then the outline for the proof.

- First, I will derive an expression for capital growth, and show how it depends on μ_τ, μ_L.
- Then, for any $M' > M$, I will consider two cases: either $\mu_L' \geq \mu_L$ or $\mu_L' < \mu_L$. I demonstrate that $\mu_L' \geq \mu_L$ results in a contradiction and therefore $\mu_L' < \mu_L$ holds. Intuitively, $\mu_L' \geq \mu_L$ leads to a contradiction, because it implies a higher market share for a_L-type firms, while at the same time increasing G_τ and thereby inducing higher market shares for a_H-type firms as well. Increasing market shares for both a_L, a_H-type firms then contradicts with the average market share decreasing with M.
- I then show that $\mu_L' < \mu_L$ implies that

- $\forall \tau > 0 : G'_\tau \leq G_\tau \wedge ((\Phi_\tau > 0) \implies G'_\tau < G_\tau)$
- $\forall \tau > 0 : \mu'_\tau \leq \mu_\tau$
- $\forall \tau > 0 : \frac{\mu'_L}{\mu'_\tau} > \frac{\mu_L}{\mu_\tau}$

which concludes the proof. This pattern for markups and capital growth is intuitive: increased M lowers markups for all types of firms, which at the same time reduces capital growth for financially constrained firms. The theoretical challenge lies in demonstrating that this is the only possible pattern for markups and capital growth.

Expression for capital growth

I consider capital growth for all firms with $\tau \geq 0$. This type of firms are only heteregenous across different bins τ, and are perfectly homogeneous within a bin τ. This will be reflected in the notation. At the same time, as explained in the paper, capital k_τ for these firms is predetermined, and their productivity a_τ is exogenous: $a_0 = a_L, \forall \tau > 0 : a_\tau = a_H$.

Financially constrained firms, i.e. firms with $\Phi_\tau > 0$, invest all their retained earnings into capital investment. Therefore, for a financially constrained firm in bin τ, capital growth $g(k_\tau) = \frac{(\mu_\tau - \frac{AC_\tau}{MC_\tau}) y_\tau MC_\tau}{k_\tau} - \delta$, where AC_τ is average cost and MC_τ is marginal cost.

The firm's total costs, for any quantity $\bar{y}_\tau$ are $TC(\bar{y}_\tau) = \frac{w}{P} L(\bar{y}_\tau)$. Here, since a_τ, and k_τ are exogenous and predetermined, respectively, setting $\bar{y}_\tau$ directly implies setting $\bar{l}_\tau$ since $\bar{y}_\tau = a_H k_\tau^\alpha \bar{l}_\tau^{1-\alpha}$. This means that $L(\bar{y}_\tau) = \left(\frac{\bar{y}_\tau}{a_H k_\tau^\alpha}\right)^{\frac{1}{1-\alpha}}$, such that $TC(\bar{y}_\tau) = \frac{w}{P}\left(\frac{\bar{y}_\tau}{a_H k_\tau^\alpha}\right)^{\frac{1}{1-\alpha}}$

Therefore:

$$MC_\tau(\bar{y}_\tau) = \frac{\partial TC(\bar{y}_\tau)}{\partial \bar{y}_\tau} = \frac{w}{(1-\alpha)P}\left(\frac{\bar{y}_\tau^\alpha}{a_H k_\tau^\alpha}\right)^{\frac{1}{1-\alpha}}$$

$$AC_\tau(\bar{y}_\tau) = \frac{w}{P}\frac{1}{\bar{y}_\tau}\left(\frac{\bar{y}_\tau}{a_H k_\tau^\alpha}\right)^{\frac{1}{1-\alpha}} = \frac{w}{P}\left(\frac{\bar{y}_\tau^\alpha}{a_H k_\tau^\alpha}\right)^{\frac{1}{1-\alpha}}$$

Which implies that

$$\frac{AC_\tau(\bar{y}_\tau)}{MC_\tau(\bar{y}_\tau)} = (1-\alpha)$$

Capital growth expression Start with derivation of profits, where $\bar{\mu}_\tau$ is determined by choosing $\bar{y}_\tau$ and setting the price given the demand function.

$$\pi_\tau = \left(\bar{\mu}_\tau - \frac{AC_\tau}{MC_\tau}\right)\bar{y}_\tau * MC_\tau = (\bar{\mu}_\tau - (1-\alpha))\frac{w}{(1-\alpha)P}\left(\frac{\bar{y}_\tau^\alpha}{a_H k_\tau^\alpha}\right)^{\frac{1}{1-\alpha}}\bar{y}_\tau$$

$$\pi_\tau = (\bar{\mu}_\tau - (1-\alpha))\frac{w}{P(1-\alpha)}\left(\frac{\bar{y}_\tau}{a_H k_\tau^\alpha}\right)^{\frac{1}{1-\alpha}}$$

Hence,

$$\frac{\pi_\tau}{k_\tau} = (\bar{\mu}_\tau - (1-\alpha)) \frac{w}{P(1-\alpha)k_\tau} \left(\frac{\bar{y}_\tau}{a_H k_\tau^\alpha}\right)^{\frac{1}{1-\alpha}} = [\bar{\mu}_\tau - (1-\alpha)] \frac{w}{P(1-\alpha)} \left(\frac{\bar{y}_\tau}{a_H k_\tau}\right)^{\frac{1}{1-\alpha}}$$

and since $\bar{y}_\tau = a_\tau k_\tau^\alpha \bar{l}_\tau^{1-\alpha}$

$$\frac{\pi_\tau}{k_\tau} = (\bar{\mu}_\tau - (1-\alpha)) \frac{w}{P(1-\alpha)} \left(\frac{a_\tau \bar{l}_\tau^{1-\alpha}}{a_\tau k_\tau^{1-\alpha}}\right)^{\frac{1}{1-\alpha}} = [\bar{\mu}_\tau - (1-\alpha)] \frac{w}{P(1-\alpha)} \frac{\bar{l}_\tau}{k_\tau}$$

Since all profits are invested in capital growth, we have for $\Phi_\tau > 0$:

$$\frac{k_{\tau+1}(\bar{l}_\tau) - k_\tau}{k_\tau} = \frac{\pi_\tau}{k_\tau} - \delta = [\bar{\mu}_\tau - (1-\alpha)] \frac{w}{P(1-\alpha)} \frac{\bar{l}_\tau}{k_\tau} - \delta$$

Given the expression for $\frac{\pi_\tau}{k_\tau}$, we need to determine $\mu_\tau, \frac{\bar{l}_\tau}{k_\tau}$. These variables are outcomes of the optimization problem, where optimal labor l_τ is from equation (1.18) in the paper, while for capital $k_\tau = G_\tau k_L$. Finally, the markup is also optimally determined as μ_τ, defined in equation (1.24).

Remember:

$$l_\tau = \left(\frac{(1-\alpha)}{\mu_\tau} \frac{P}{w} \left(\frac{Q}{M}\right)^{1-\eta} a_\tau^\eta k_\tau^{\alpha\eta}\right)^{\frac{1}{1+\alpha\eta-\eta}}$$

$$k_L^* = \left(\frac{a_L^\eta}{\mu_L}\right)^{\frac{1}{1-\eta}} \left(\frac{\alpha}{r_L}\right)^{\frac{1+\alpha\eta-\eta}{1-\eta}} \left(\frac{P(1-\alpha)}{w}\right)^{\frac{\eta-\alpha\eta}{1-\eta}} \frac{Q}{M}$$

therefore

$$\frac{l_\tau^*}{G_\tau k_L^*} = \frac{P(1-\alpha)}{G_\tau w} \left(\frac{a_\tau}{a_L}\right)^{\frac{\eta}{1+\alpha\eta-\eta}} \frac{r_L}{\alpha} \left(\frac{\mu_L}{\mu_\tau}\right)^{\frac{1}{1+\alpha\eta-\eta}}$$

which implies that for $g(k)_\tau \equiv \frac{k_{\tau+1}(l_\tau)-k_\tau}{k_\tau}$

$$\forall \tau \text{ where } \Phi_\tau > 0 : g(k)_\tau = \frac{1}{G_\tau} [\mu_\tau - (1-\alpha)] \left(\frac{a_\tau}{a_L}\right)^{\frac{\eta}{1+\alpha\eta-\eta}} \frac{r_L}{\alpha} \left(\frac{\mu_L}{\mu_\tau}\right)^{\frac{1}{1+\alpha\eta-\eta}} - \delta$$

or

$$\forall \tau \text{ with } \Phi_\tau > 0 : \ln(g(k)_\tau G_\tau + \delta) = \ln [\mu_\tau - (1-\alpha)] + \ln \left(\frac{a_\tau}{a_L}\right)^{\frac{\eta}{1+\alpha\eta-\eta}} \frac{r_L}{\alpha} - \frac{\ln\left(\frac{\mu_\tau}{\mu_L}\right)}{1+\alpha\eta-\eta} \tag{1.33}$$

Derivative with respect to M Assuming r_L is constant, and only considering τ where $\Phi_\tau > 0$ we find that

$$\frac{\partial \ln(g(k)_\tau G_\tau + \delta)}{\partial M} = \frac{\frac{\partial \mu_\tau}{\partial M}}{[\mu_\tau - (1-\alpha)]} - \frac{1}{1+\alpha\eta-\eta}\frac{\mu_L}{\mu_\tau}\left[\frac{\partial \mu_\tau}{\partial M}\frac{1}{\mu_L} - \frac{\mu_\tau}{\mu_L^2}\frac{\partial \mu_L}{\partial M}\right]$$

Rearranging:

$$\frac{\partial \ln(g(k)_\tau G_\tau + \delta)}{\partial M} = \frac{\partial \mu_\tau}{\partial M}\left(\frac{1}{[\mu_\tau - (1-\alpha)]} - \frac{1}{1+\alpha\eta-\eta}\frac{1}{\mu_\tau}\left[1 - \frac{\mu_\tau}{\mu_L}\frac{\frac{\partial \mu_L}{\partial M}}{\frac{\partial \mu_\tau}{\partial M}}\right]\right) \tag{1.34}$$

Note that a sufficient condition for $sign(\frac{\partial \ln(g(k)_\tau G_\tau+\delta)}{\partial M}) = sign(\frac{\partial \mu_\tau}{\partial M})$, is that $sign(\frac{\partial \mu_\tau}{\partial M}) = sign(\frac{\partial \mu_L}{\partial M})$. This is because

$$\left(\frac{1}{[\mu_\tau - (1-\alpha)]} - \frac{1}{1+\alpha\eta-\eta}\frac{1}{\mu_\tau}\left[1 - \frac{\mu_\tau}{\mu_L}\frac{\frac{\partial \mu_L}{\partial M}}{\frac{\partial \mu_\tau}{\partial M}}\right] > 0\right) \iff \left(1 > \frac{\left[1 - \frac{(1-\alpha)}{\mu_\tau}\right]}{1-\eta(1-\alpha)}\left[1 - \frac{\mu_\tau}{\mu_L}\frac{\frac{\partial \mu_L}{\partial M}}{\frac{\partial \mu_\tau}{\partial M}}\right]\right)$$

and $\left(sign(\frac{\partial \mu_\tau}{\partial M}) = sign(\frac{\partial \mu_L}{\partial M})\right) \implies \left(0 > -\frac{\mu_\tau}{\mu_L}\frac{\frac{\partial \mu_L}{\partial M}}{\frac{\partial \mu_\tau}{\partial M}}\right)$. Hence, a key step in the remainder of the proof will be demonstrating that $\left(sign(\frac{\partial \mu_\tau}{\partial M}) = sign(\frac{\partial \mu_L}{\partial M})\right)$ holds globally.

Impact of competition on distribution of markups and capital growth

Given equation (1.34), I will now examine the level of markups and capital growth, across any two different levels for the number of firms in the economy, namely $M' > M$, where I denote with a prime the values under M'. Specifically, I will examine two cases. First, $\mu_L' \geq \mu_L$ and second $\mu_L' < \mu_L$. The first case will result in a contradiction, so its opposite - the second case - must be true. The analysis in the second case will then characterize the path of markups and capital growth across different M. For the analysis, it will be useful to define $\mathcal{G}_\tau \equiv \frac{y_\tau^\eta}{y_L^\eta}$.

Case 1: $\mu_L' \geq \mu_L$

This case will result in a contradiction, and therefore its opposite must be true. The proof proceeds by induction.

Step 1 Consider $\tau = 0$, where productivity is a_L and $\mu_0 = \mu_L$, but the firm learns it will have productivity a_H in $\tau = 1$. In this period, if $\Phi_0 > 0$, capital growth is

$$g(k)_0 = [\mu_L - (1-\alpha)]\,\frac{r_L}{\alpha} \tag{1.35}$$

Therefore $g(k)'_0 \geq g(k)_0$ since $\mu'_L \geq \mu_L$ and the other variables are constant.

Inductive step For the inductive step, I show first that for any period $\tau > 0$ with $G'_\tau \geq G_\tau$:

$$((G'_\tau \geq G_\tau) \wedge (\mu'_L \geq \mu_L)) \implies (\mu'_\tau \geq \mu_\tau)$$

To show this, notice that $(\mathcal{G}'_\tau \geq \mathcal{G}_\tau)/(\mathcal{G}'_\tau < \mathcal{G}_\tau)$, and in both cases, I show that $(\mu'_\tau \geq \mu_\tau)$ holds

- Case (i): $((\mathcal{G}'_\tau \geq \mathcal{G}_\tau) \wedge (G'_\tau \geq G_\tau) \wedge (\mu'_L \geq \mu_L)) \implies (\mu'_\tau \geq \mu_\tau)$. This follows from $((\mathcal{G}'_\tau \geq \mathcal{G}_\tau) \wedge (\mu'_L \geq \mu_L)) \implies (\mu'_\tau \geq \mu_\tau)$.
  - From equations (1.9), (1.10), it is clear that μ_L is monotonically increasing in $\frac{y_L^\eta}{\sum y_{it}^\eta}$. Therefore, $(\mu'_L \geq \mu_L) \iff \left(\frac{y'^\eta_L}{\sum y'^\eta_{it}} \geq \frac{y_L^\eta}{\sum y_{it}^\eta}\right)$. Then, $(\mathcal{G}'_\tau \geq \mathcal{G}_\tau) \wedge \left(\frac{y'^\eta_L}{\sum y'^\eta_{it}} \geq \frac{y_L^\eta}{\sum y_{it}^\eta}\right) \implies \left(\frac{\mathcal{G}'_\tau y'^\eta_L}{\sum y'^\eta_{it}} \geq \frac{\mathcal{G}_\tau y_L^\eta}{\sum y_{it}^\eta} \wedge \mu'_\tau \geq \mu_\tau\right)$
- Case (ii): $((\mathcal{G}'_\tau < \mathcal{G}_\tau) \wedge (G'_\tau \geq G_\tau) \wedge (\mu'_L \geq \mu_L)) \implies (\mu'_\tau \geq \mu_\tau)$.
  - Note that $\mathcal{G}_\tau \equiv \frac{y_\tau^\eta}{y_L^\eta} = \frac{(a_H G_\tau^\alpha l_\tau^{1-\alpha})^\eta}{(a_L l_L^{1-\alpha})^\eta}$. Therefore $(\mathcal{G}'_\tau < \mathcal{G}_\tau) \implies \left((G'_\tau < G_\tau) \vee (\frac{l'_\tau}{l'_L} < \frac{l_\tau}{l_L})\right)$ such that $\left((G'_\tau \geq G_\tau) \wedge (\frac{l'_\tau}{l'_L} \geq \frac{l_\tau}{l_L})\right) \implies (\mathcal{G}'_\tau \geq \mathcal{G}_\tau)$
  - Note that $\frac{l_\tau}{l_L} = \left(\frac{\mu_L}{\mu_\tau}\frac{a_H}{a_L}G_\tau^{\alpha\eta}\right)^{\frac{1}{1+\alpha\eta-\eta}}$. Therefore, $\left((\frac{l'_\tau}{l'_L} < \frac{l_\tau}{l_L}) \wedge (G'_\tau \geq G_\tau)\right) \implies \frac{\mu'_L}{\mu'_\tau} < \frac{\mu_L}{\mu_\tau}$
  - In this case, $(\mathcal{G}'_\tau < \mathcal{G}_\tau) \wedge (G'_\tau \geq G_\tau) \implies \left(\frac{l'_\tau}{l'_L} < \frac{l_\tau}{l_L}\right)$. However, $\left(\frac{l'_\tau}{l'_L} < \frac{l_\tau}{l_L}\right) \implies \frac{\mu'_L}{\mu'_\tau} < \frac{\mu_L}{\mu_\tau}$. Therefore, $((\mathcal{G}'_\tau < \mathcal{G}_\tau) \wedge (G'_\tau \geq G_\tau) \wedge (\mu'_L \geq \mu_L)) \implies (\mu'_\tau > \mu_\tau)$.
- Therefore, $((G'_\tau \geq G_\tau) \wedge (\mu'_L \geq \mu_L)) \implies (\mu'_\tau \geq \mu_\tau)$,

Given equation (1.34), $((\Phi_\tau > 0) \wedge (\mu'_\tau > \mu_\tau) \wedge (\mu'_L \geq \mu_L) \wedge (G'_\tau \geq G_\tau)) \implies G'_{\tau+1} \geq G_{\tau+1}$. This completes the inductive step.

Final step In case $\Phi_0 > 0$, then $g(k)'_0 \geq g(k)_0$. Hence, $(G'_1 \geq G_1) \implies (\mu'_1 \geq \mu_1)$. Therefore, this proof by induction implies that $(\mu'_L \geq \mu_L) \implies (\forall \tau > 0 \wedge \Phi_\tau > 0 : (\mu'_\tau \geq \mu_\tau)$. At the same time, $((M' > M) \wedge (\mu'_L > \mu_L)) \implies \exists \tau > 0 : (\mu'_\tau < \mu_\tau)$, which will yield a contradiction.

- To see that $((M' > M) \wedge (\mu'_L > \mu_L)) \implies \exists\tau : (\mu'_\tau < \mu_\tau)$, note that equations (1.9) and (1.10) for the markup and the demand elasticity entail that for any firm i: $(\mu'_{it} \geq \mu_{it}) \iff \left(\frac{y'^{\eta}_{it}}{\sum_i y'^{\eta}_{it}} \geq \frac{y^{\eta}_{it}}{\sum_i y^{\eta}_{it}}\right)$. Suppose for $M' > M$, we have $((\mu'_L \geq \mu_L) \wedge (\forall\tau > 0 : \mu'_\tau \geq \mu_\tau)) \implies \left((\frac{y'^{\eta}_L}{\sum_i y'^{\eta}_{it}} \geq \frac{y^{\eta}_L}{\sum_i y^{\eta}_{it}}) \wedge (\frac{y'^{\eta}_\tau}{\sum_i y'^{\eta}_{it}} \geq \frac{y^{\eta}_\tau}{\sum_i y^{\eta}_{it}})\right)$. Then,

$$\left((\frac{y'^{\eta}_L}{\sum_i y'^{\eta}_{it}} \geq \frac{y^{\eta}_L}{\sum_i y^{\eta}_{it}}) \wedge (\frac{y'^{\eta}_\tau}{\sum_i y'^{\eta}_{it}} \geq \frac{y^{\eta}_\tau}{\sum_i y^{\eta}_{it}})\right)$$
$$\implies \left(\frac{M'[Prob(a_{it} = a_L)\frac{y'^{\eta}_L}{\sum_i y'^{\eta}_{it}} + \sum_{\tau=1}^{\infty} Prob((a_{it} = a_H)\&(t = \tau))\frac{y'^{\eta}_\tau}{\sum_i y'^{\eta}_{it}}]}{M[Prob(a_{it} = a_L)\frac{y^{\eta}_L}{\sum_i y^{\eta}_{it}} + \sum_{\tau=1}^{\infty} Prob((a_{it} = a_H)\&(t = \tau))\frac{y^{\eta}_\tau}{\sum_i y^{\eta}_{it}}]} > 1\right)$$

  Which is a contradiction since both the denominator and the numerator in the ratio after the implication are equal to 1. This is because $\sum_i y^{\eta}_{it} = M[Prob(a_{it} = a_L)y^{\eta}_L + \sum_{\tau=1}^{\infty} Prob((a_{it} = a_H)\&(t = \tau))y^{\eta}_\tau]$. Since $M' > M \wedge ((\mu'_L \geq \mu_L) \wedge (\forall\tau > 0 : \mu'_\tau \geq \mu_\tau))$ entails a contradiction, $\exists\tau > 0 : (\mu'_\tau < \mu_\tau)$, under the continued assumption that $\mu'_L \geq \mu_L$.

- $(\exists\tau > 0 : (\mu'_\tau < \mu_\tau)) \implies ((\exists\tau > 0 : \Phi_\tau > 0 \wedge (\mu'_\tau < \mu_\tau)) \vee (\exists\tau > 0 : \Phi_\tau = 0 \wedge (\mu'_\tau < \mu_\tau)))$, but both cases result in a contradiction.
  - Case a: $(\exists\tau > 0 : \Phi_\tau > 0 \wedge (\mu'_\tau < \mu_\tau))$. This does not hold, since the proof by induction implies $(\mu'_L \geq \mu_L) \implies (\forall\tau > 0 \wedge \Phi_\tau > 0 : (\mu'_\tau \geq \mu_\tau)$.
  - Case b is equivalent to $(\mu'_H < \mu_H)$. We know that $(\mu'_L > \mu_L) \wedge (\mu'_H < \mu_H) \implies (\mathcal{G}'_H < \mathcal{G}_H)$, where $\mathcal{G}_H = \frac{(a_H G_H^{\alpha} l_H^{1-\alpha})^{\eta}}{a_L l_L^{1-\alpha})^{\eta}}$. Hence, $(\mathcal{G}'_H < \mathcal{G}_H) \implies \left((G'_H < G_H) \vee (\frac{l'_H}{l'_L} < \frac{l_H}{l_L})\right)$. There are then again two cases, both of which result in a contradiction:
    - Case b1: since $G_H = (\frac{a_H}{a_L}\frac{\mu_L}{\mu_H})^{1/(1-\eta)}$, $(G'_H < G_H) \implies (\frac{\mu'_L}{\mu'_H} < \frac{\mu_L}{\mu_H})$. However, $(\mu'_L > \mu) \wedge (\frac{\mu'_L}{\mu'_H} < \frac{\mu_L}{\mu_H}) \implies (\mu'_H > \mu_H)$, which contradicts the supposition that $(\mu'_H < \mu_H)$
    - Case b2: Since $\frac{l_\tau}{l_L} = \left(\frac{\mu_L}{\mu_\tau}\frac{a_H}{a_L}G_\tau^{\alpha\eta}\right)^{\frac{1}{1+\alpha\eta-\eta}}$, $(\frac{l'_H}{l'_L} < \frac{l_H}{l_L}) \wedge (G'_H \geq G_H) \implies (\frac{\mu'_L}{\mu'_H} < \frac{\mu_L}{\mu_H})$, which again results in a contradiction

Since the supposition that $\mu'_L \geq \mu_L$ entails a contradiction, its opposite must be true: $\mu'_L < \mu_L$

Case 2: $\mu'_L < \mu_L$

Step 1 Consider $\tau = 0$, from equation (1.35), it is clear that

$$((\mu'_L < \mu_L) \wedge (\Phi_0 > 0)) \implies ((g(k)'_0 < g(k)_0) \wedge (G'_1 < G_1))$$

Inductive step For the inductive step, I show first that for any period $\tau > 0$:

$$(G'_\tau < G_\tau) \wedge (\mu'_L < \mu_L) \implies (\mu'_\tau < \mu_\tau)$$

To prove that $(G'_\tau < G_\tau) \wedge (\mu'_L < \mu_L) \implies (\mu'_\tau < \mu_\tau)$, consider two cases:

- Case (i): $(G'_\tau < G_\tau) \wedge (\mu'_L < \mu_L) \wedge (\mathcal{G}'_\tau \leq \mathcal{G}_\tau) \implies (\mu'_\tau < \mu_\tau)$.
- Case (ii): $(G'_\tau < G_\tau) \wedge (\mu'_L < \mu_L) \wedge (\mathcal{G}'_\tau > \mathcal{G}_\tau) \implies (\mu'_\tau < \mu_\tau)$. This is because given $\mathcal{G}_\tau = \frac{(a_H G_\tau l_\tau)^\eta}{(a_L l_L)^\eta}$ and $\frac{l_\tau}{l_L} = \left(\frac{\mu_L}{\mu_\tau}\frac{a_H}{a_L}G_\tau^{\alpha\eta}\right)^{\frac{1}{1+\alpha\eta-\eta}}$; $((G'_\tau < G_\tau) \wedge (\mathcal{G}'_\tau > \mathcal{G}_\tau)) \implies (\frac{l'_1}{l'_L} > \frac{l_1}{l_L})$ and $\left((\mu'_L < \mu_L) \wedge (\frac{l'_1}{l'_L} > \frac{l_1}{l_L}) \wedge (G'_\tau < G_\tau)\right) \implies \left((\frac{\mu'_L}{\mu'_\tau} > \frac{\mu_L}{\mu_\tau}) \iff (1 > \frac{\mu'_L}{\mu_L} > \frac{\mu'_\tau}{\mu_\tau})\right)$
- Therefore, $(G'_\tau < G_\tau) \wedge (\mu'_L < \mu_L) \implies (\mu'_\tau < \mu_\tau)$

Given equation (1.34), $((\Phi_\tau > 0) \wedge (\mu'_\tau < \mu_\tau) \wedge (\mu'_L < \mu_L) \wedge (G'_\tau < G_\tau)) \implies G'_{\tau+1} < G_{\tau+1}$. This completes the inductive step, which applies for any $\tau > 0$ with $\Phi_\tau > 0$.

Final step We know that when $\Phi_0 > 0$, $g(k)'_0 < g(k)_0$. Hence, $\Phi_0 > 0 \implies [(G'_1 < G_1) \implies (\mu'_1 < \mu_1)]$. Therefore, this proof by induction implies that

$$(\mu'_L < \mu_L) \implies [(\Phi_\tau > 0) \implies ((\mu'_\tau < \mu_\tau) \wedge (G'_\tau < G_\tau))]$$

Result for μ_H, G_H How do μ_H, G_H evolve with M? There are two cases: $\mathcal{G}'_H \leq \mathcal{G}_H$ or $\mathcal{G}'_H > \mathcal{G}_H$.

- Case a: $(\mathcal{G}'_H \leq \mathcal{G}_H) \implies (\mu'_H < \mu_H)$. Why? We know that $\mu'_L < \mu_L \iff (\frac{y'^\eta_L}{\sum_i y'^\eta_{it}} < \frac{y^\eta_L}{\sum_i y^\eta_{it}})$. Hence, $((\mu'_L < \mu_L) \wedge (\mathcal{G}'_H \leq \mathcal{G}_H)) \implies \left((\mathcal{G}'_H \frac{y'^\eta_L}{\sum_i y'^\eta_{it}} < \frac{\mathcal{G}_H y^\eta_L}{\sum_i y^\eta_{it}}) \iff (\mu'_H < \mu_H)\right)$
- Case b: $(\mathcal{G}'_H > \mathcal{G}_H)$.
  - Note that $\mathcal{G}_\tau \equiv \frac{y^\eta_H}{y^\eta_L} = \frac{(a_H G_\tau l_H)^\eta}{(a_L l_L^{1-\alpha})^\eta}$. Hence, $(\mathcal{G}'_H > \mathcal{G}_H) \iff (\frac{G'^\alpha_H l'^{1-\alpha}_H}{l'_L} < \frac{G^\alpha_H l^{1-\alpha}_H}{l_L})$. Therefore $(\mathcal{G}'_H > \mathcal{G}_H) \implies \left((G'_H > G_H) \vee (\frac{l'_H}{l'_L} > \frac{l_H}{l_L})\right)$. There are then again two cases
  - Case b1: suppose $(G'_H > G_H)$. Since $G_H = (\frac{a_H}{a_L}\frac{\mu_L}{\mu_H})^{1/(1-\eta)}$, $(G'_H > G_H) \implies \left((\frac{\mu'_L}{\mu'_H} > \frac{\mu_L}{\mu_H}) \iff (\frac{\mu'_L}{\mu_L} > \frac{\mu'_H}{\mu_H})\right)$. Therefore, $(G'_H > G_H) \wedge (\mu'_L < \mu_L) \implies$
  - Case b2: suppose $(\frac{l'_H}{l'_L} > \frac{l_H}{l_L}) \wedge (G'_H \leq G_H)$. Note that $\frac{l_H}{l_L} = \left(\frac{\mu_L}{\mu_H}\frac{a_H}{a_L}G_H^{\alpha\eta}\right)^{\frac{1}{1+\alpha\eta-\eta}}$. Therefore, $\left((\frac{l'_H}{l'_L} > \frac{l_H}{l_L}) \wedge (G'_H \leq G_H)\right) \implies (\frac{\mu'_L}{\mu'_H} > \frac{\mu_L}{\mu_H} \iff \frac{\mu'_L}{\mu_L} > \frac{\mu'_H}{\mu_H})$. Since $1 > \frac{\mu'_L}{\mu_L}$, we find that $(\frac{l'_H}{l'_L} > \frac{l_H}{l_L}) \wedge (G'_H \leq G_H) \implies (\mu'_H < \mu_H)$.

– Therefore, we find that $(\mu'_H < \mu_H)$ and hence

$$(\mu'_L < \mu_L) \implies [\forall \tau > 0 : ((\mu'_\tau < \mu_\tau))]$$

Relative markups and M

From the previous subsection, I know that

$$(\mu'_L < \mu_L) \wedge (\forall \tau > 0 : \mu'_\tau < \mu_\tau) \wedge ((\Phi_\tau > 0) \implies (G'_\tau < G_\tau))$$

As is already clear from the inductive step in subsection 1.3, there are two cases: either $\mathcal{G}'_\tau \leq \mathcal{G}_\tau$ or $\mathcal{G}'_\tau > \mathcal{G}_\tau$. I now demonstrate that in both cases, $\frac{\mu'_L}{\mu'_\tau} > \frac{\mu_L}{\mu_\tau}$.

- In case $\mathcal{G}'_\tau > \mathcal{G}_\tau$, then case (ii) in subsection 1.3 demonstrates that

$$((G'_\tau < G_\tau) \wedge (\mu'_L < \mu_L) \wedge (\mathcal{G}'_\tau > \mathcal{G}_\tau)) \implies \left(\frac{\mu'_L}{\mu'_\tau} > \frac{\mu_L}{\mu_\tau}\right)$$

- In case $\mathcal{G}'_\tau \leq \mathcal{G}_\tau$, then start from the expression for relative markups, derived from equation (1.9):

$$\frac{\mu_L}{\mu_\tau} = \frac{\frac{1-\eta\frac{y_L^\eta}{\sum_i y_{it}^\eta}}{\eta\left(1-\frac{y_L^\eta}{\sum_i y_{it}^\eta}\right)}}{\frac{1-\eta\frac{y_\tau^\eta}{\sum_i y_{it}^\eta}}{\eta\left(1-\frac{y_\tau^\eta}{\sum_i y_{it}^\eta}\right)}} = \frac{1-\eta\frac{y_L^\eta}{\sum_i y_{it}^\eta}}{\left(1-\frac{y_L^\eta}{\sum_i y_{it}^\eta}\right)} \frac{\left(1-\frac{y_\tau^\eta}{\sum_i y_{it}^\eta}\right)}{1-\eta\frac{y_\tau^\eta}{\sum_i y_{it}^\eta}}$$

First, define $\frac{y_\tau^\eta}{y_L^\eta} \equiv \mathcal{G}_\tau$, such that:

$$\frac{\mu_L}{\mu_\tau} = \frac{\left(1-\eta\frac{y_L^\eta}{\sum_i y_{it}^\eta}\right)}{\left(1-\frac{y_L^\eta}{\sum_i y_{it}^\eta}\right)} \frac{\left(1-\frac{\mathcal{G}_\tau y_L^\eta}{\sum_i y_{it}^\eta}\right)}{\left(1-\eta\frac{\mathcal{G}_\tau y_L^\eta}{\sum_i y_{it}^\eta}\right)} = \frac{1-\frac{y_L^\eta}{\sum_i y_{it}^\eta}(\mathcal{G}_\tau+\eta)+\eta\mathcal{G}_\tau(\frac{y_L^\eta}{\sum_i y_{it}^\eta})^2}{1-\frac{y_L^\eta}{\sum_i y_{it}^\eta}(\mathcal{G}_\tau\eta+1)+\eta\mathcal{G}_\tau(\frac{y_L^\eta}{\sum_i y_{it}^\eta})^2}$$

Define: $Num \equiv 1-\frac{y_L^\eta}{\sum_i y_{it}^\eta}(\mathcal{G}_\tau+\eta)+\eta\mathcal{G}_\tau(\frac{y_L^\eta}{\sum_i y_{it}^\eta})^2$ and $Denom \equiv 1-\frac{y_L^\eta}{\sum_i y_{it}^\eta}(\mathcal{G}_\tau\eta+1)+\eta\mathcal{G}_\tau(\frac{y_L^\eta}{\sum_i y_{it}^\eta})^2$ and find that

$$\begin{aligned}
&\frac{\partial\frac{\mu_L}{\mu_\tau}}{\partial M} * Denom^2 = \\
&\left[-\frac{\partial\frac{y_L^\eta}{\sum_i y_{it}^\eta}}{\partial M}(\mathcal{G}_\tau+\eta) - \frac{y_L^\eta}{\sum_i y_{it}^\eta}\frac{\partial\mathcal{G}_\tau}{\partial M} + 2\mathcal{G}_\tau\eta\frac{y_L^\eta}{\sum_i y_{it}^\eta}\frac{\partial\frac{y_L^\eta}{\sum_i y_{it}^\eta}}{\partial M} + \eta(\frac{y_L^\eta}{\sum_i y_{it}^\eta})^2\frac{\partial\mathcal{G}_\tau}{\partial M}\right] * Denom \\
&- Num\left[-\frac{\partial\frac{y_L^\eta}{\sum_i y_{it}^\eta}}{\partial M}(\mathcal{G}_\tau\eta+1) - \frac{y_L^\eta}{\sum_i y_{it}^\eta}\eta\frac{\partial\mathcal{G}_\tau}{\partial M} + 2\eta\mathcal{G}_\tau\frac{\partial\frac{y_L^\eta}{\sum_i y_{it}^\eta}}{\partial M} + \eta(\frac{y_L^\eta}{\sum_i y_{it}^\eta})^2\frac{\partial\mathcal{G}_\tau}{\partial M}\right]
\end{aligned}$$

Rearranging the RHS:

$$\frac{\partial \frac{y_L^\eta}{\sum_i y_{it}^\eta}}{\partial M}\left[Num(\mathcal{G}_\tau\eta+1)-Denom(\mathcal{G}_\tau+\eta)+2\frac{y_L^\eta}{\sum_i y_{it}^\eta}\mathcal{G}_\tau\eta(Denom-Num)\right]$$
$$+\eta\frac{y_L^\eta}{\sum_i y_{it}^\eta}\frac{\partial\mathcal{G}_\tau}{\partial M}\left[(Num-\frac{Denom}{\eta})+\frac{y_L^\eta}{\sum_i y_{it}^\eta}(Denom-Num)\right]$$

Note that $\left[(Num-\frac{Denom}{\eta})+\frac{y_L^\eta}{\sum_i y_{it}^\eta}(Denom-Num)\right] < 0$ for any $\frac{y_L^\eta}{\sum_i y_{it}^\eta} < 1$ since $\eta < 1$ and $Denom > Num$. Since $\frac{\partial \frac{y_L^\eta}{\sum_i y_{it}^\eta}}{\partial M} < 0$ because $\mu_L' < \mu_L$ and because I assume that $\frac{\partial\mathcal{G}_\tau}{\partial M} < 0$, the following condition is sufficient for $\frac{\partial \frac{\mu_L}{\mu_\tau}}{\partial M} > 0$ to hold:

$$\left[Num(\mathcal{G}_\tau\eta+1)-Denom(\mathcal{G}_\tau+\eta)+2\mathcal{G}_\tau\eta\frac{y_L^\eta}{\sum_i y_{it}^\eta}(Denom-Num)\right]<0$$

or

$$2\mathcal{G}_\tau\eta\frac{y_L^\eta}{\sum_i y_{it}^\eta}(Denom-Num) < Denom(\mathcal{G}_\tau+\eta)-Num(\mathcal{G}_\tau\eta+1)$$

$$\frac{y_L^\eta}{\sum_i y_{it}^\eta} < \frac{Denom(\mathcal{G}_\tau+\eta)-Num(\mathcal{G}_\tau\eta+1)}{2\mathcal{G}_\tau\eta(Denom-Num)} = \frac{\mathcal{G}_\tau+\eta}{2\mathcal{G}_\tau\eta}-\frac{Num(\eta-1)(\mathcal{G}_\tau-1)}{2\mathcal{G}_\tau\eta(Denom-Num)}$$

Hence, the following condition is more than sufficient for $\frac{\partial \frac{\mu_L}{\mu_\tau}}{\partial M} > 0$ to hold:

$$\frac{y_L^\eta}{\sum_i y_{it}^\eta} < \frac{\mathcal{G}_\tau+\eta}{2\mathcal{G}_\tau\eta}+\frac{Num(1-\eta)(\mathcal{G}_\tau-1)}{2\mathcal{G}_\tau\eta(Denom-Num)}$$

If we only consider cases with $M > 2$, then $\frac{y_L^\eta}{\sum_i y_{it}^\eta} < \frac{1}{2}$, and a sufficient condition is:

$$1 < \frac{\mathcal{G}_\tau+\eta}{\mathcal{G}_\tau\eta}+\frac{Num(1-\eta)(\mathcal{G}_\tau-1)}{\mathcal{G}_\tau\eta(Denom-Num)}$$

This holds for any value $0 \leq \eta \leq 1$, since it holds for $\eta = 0, 1$ and for $\eta > 0$, the RHS is monotonically declining because $\frac{\partial \frac{\mathcal{G}_\tau+\eta}{\mathcal{G}_\tau\eta}}{\partial\eta} = \frac{\mathcal{G}_\tau\eta-(\mathcal{G}_\tau+\eta)\mathcal{G}}{(\mathcal{G}_\tau\eta)^2} = -\frac{1}{\eta^2}$. Hence, we always have that $\frac{\partial \frac{\mu_L}{\mu_\tau}}{\partial M} > 0$.

1.4 Concluding Remarks

This chapter examines the theoretical relation between capital misallocation and the degree of competition. The theory describes how competition affects steady-state misallocation in a setting with firm-level productivity volatility and financial constraints. Competition plays a dual role in affecting misallocation. On the one hand, both markup levels and markup dispersion tend to fall with competition, which unambiguously improves allocative efficiency in a setting without financial constraints. On the other hand, in a setting with financial constraints, a reduction in markups slows down capital accumulation, as the rate of self-financed investment shrinks. Thus, the positive impact of competition on steady-state misallocation is reduced by the presence of financial frictions. While the beneficial impact of competition is well known in the misallocation literature, the negative impact of competition in a setting with financial constraints was previously underexamined.

Chapter 2

Competition and Capital Convergence: Plant-level Evidence from Indian Manufacturing

2.1 Introduction

In the first chapter of this dissertation, I developed a model where competition has a negative impact on capital misallocation. In this empirical chapter, I focus on testing the predictions of the model in the context of the Indian manufacturing sector. I first test the main mechanism of the model, namely that firm-level speed of capital convergence decreases with competition. This prediction can be tested at two levels: for firms in general and for young plants in particular. For firms in general, I test whether, after a firm deviates from its optimal marginal revenue product of capital (MRPK), it converges back faster to its optimal MRPK in a setting with less competition. Then, based on the model's prediction for undercapitalized young plants, I check if the capital growth rate of young plants is faster in settings where competition is less intense. These two empirical tests are complementary. The first test is closely linked to the structure of the model as it focuses directly on plant-level MRPK, where, inspired by Asker, Collard-Wexler, and De Loecker (2014), plant-level deviations in MRPK serve as a proxy of the plant-level capital wedges. The second test, which focuses on young plants, has the advantage that capital-growth is a reduced-form object in the data, and therefore relies on fewer assumptions for its measurement. The fact that both tests empirically validate the model predictions, therefore provides robust support for the model.

A second set of tests leverages heterogeneity in firms' financial dependence, where capital convergence of firms in sectors with higher financial dependence exhibits a stronger sensitivity to the degree of competition. To test this prediction, I augment the baseline tests with an interaction term of the competition measure with Rajan and Zingales (1998) measures of

sector-level financial dependence. The data again support the predictions, both for the test on MRPK convergence, and for the test on capital growth for young firms.

These two sets of predictions rely on a measure for competition that is arguably exogenous from the firm's point of view, namely the median markup measured at the state-sector level. The advantage of this approach is that I can test the theory on a large subset of the Indian manufacturing sector, while a potential limitation is that the underlying structural drivers of the variation in state-sector levels of competition remain unexamined. To address this concern, I also exploit natural variation in the degree of competition arising from India's 1997 dereservation reform, and now test whether convergence of MRPK is slower after dereservation, and whether young plants grow capital more slowly after dereservation.[1] The data again confirm the two predictions of the model.

In addition to testing the theoretical model from chapter one, the current chapter also contributes to our understanding of high and persistent level of misallocation in the Indian manufacturing sector. The empirical literature on this topic was ignited by Hsieh and Klenow (2009), who argue that misallocation could account for 40 to 60% of the difference in aggregate output per capita between the United States and India. In related work, Bollard, Klenow, and Sharma (2013) document that the persistence of misallocation is related to a broad lack of reallocation. These findings are related to the evidence on firm stagnation in India by (Hsieh and Klenow, 2014). This chapter then explains how increased competition, which is potentially associated with India's liberalization policies, can contribute to this high and persistent misallocation. As indicated above, the adverse effect of competition depends, amongst others, on the degree of productivity volatility and the entry-rate of newborn firms in a context of financial constraints. The stylized facts indicate that all these factors are substantially present in Indian manufacturing. First, for productivity volatility, Asker, Collard-Wexler, and De Loecker (2014) demonstrate that there is a strong correlation between productivity volatility and their measure of capital misallocation in the case of India. Second, Bollard, Klenow, and Sharma (2013) document high entry-rates of new firms in Indian manufacturing. Third, Banerjee and Duflo (2014) estimate severe credit constraints for large Indian firms, which is consistent with the descriptive evidence on financial constraints from the World Bank Enterprise Surveys (Kuntchev et al., 2014). Together, these three stylized facts on productivity volatility, arrival rate of newborn firms, and financial constraints, indicate the relevance of this paper for understanding misallocation in Indian manufacturing.

[1] The dereservation reform gradually removed previously existing investment ceilings on a set of "reserved" products (García-Santana and Pijoan-Mas, 2014; Martin, Nataraj, and Harrison, 2014; Tewari and Wilde, 2014). Hence, the direct effect of dereservation is to allow incumbent firms to increase their capital stock. However, the reform also leads to intensified competition, e.g. through larger firms starting to produce the previously reserved products, and this competitive channel empirically dominates in my analysis of capital convergence.

2.2 Data

The empirical analysis employs plant-level panel data from the Indian Annual Survey of Industries (ASI), for the period 1990-2011. The ASI sampling scheme consists of two components.[2] One component is a census of all manufacturing establishments with more than 100 employees, while a second component samples, with a certain probability, each formally registered establishment with less than 100 employees. All establishments with more than 20 workers (10 workers if the establishment uses electricity) are required to be formally registered.[3]

In the empirical exercise, I will be exploiting variation across sectors and geographical units in India. Here, sectors are defined as 3-digit sectors based on India's 1987 National Industrial Classification (NIC). The geographical units in the data are either states or union territories. For convenience, I will be referring to both geographical units as "states."[4]

Variable definitions

The main plant-level variables are capital K_{irst}, labor L_{irst}, materials M_{irst} and revenue S_{irst}, for plant i, state r, sector s and year t. Here, t stands for the financial year, and K_{irst} is the book value of assets at the start of the financial year. The logarithm of a variable will be denoted in lower case.

The empirical analysis will provide both motivating macro-level stylized facts, as well as micro-level evidence on capital convergence. Both sections of the empirical discussion will examine data-patterns related to plant-level capital growth, marginal revenue product of capital and markups. I now describe the construction of these main variables. First, capital growth is measured as:[5]

$$g(k_{irst}) = k_{irst+1} - k_{irst}$$

Second, marginal revenue product of capital (MRPK) is measured as in Asker, Collard-Wexler, and De Loecker (2014), who assume a sector-level Cobb-Douglas production function, which implies that the marginal revenue product of capital takes the following form:

$$MRPK_{irst} = \ln(\beta_s^K) + s_{irst} - k_{irst} \tag{2.1}$$

[2]The particulars provided here hold for the majority of the sample years. Bollard, Klenow, and Sharma (2013) provide a more detailed description of the ASI data, including certain modifications to the sampling scheme.

[3]For the years 1998-2011, establishment identifiers are provided by the Indian Statistical Office. For the pre-1998 years, I use the panel-identifiers employed by Allcott, Collard-Wexler, and O'Connell (2014), which were generously made available by Hunt Allcott.

[4]To make the definitions of states consistent over time, I employ the concordance provided by the Indian Statistical Office. This results in a number of 35 states in the panel data.

[5]Here, K_{irst+1} is the book value of assets at the end of the financial year.

Here, I also employ a value-added measure for MRPK as a leading robustness check, where $MRPK^{VA}_{irst} = \ln(\beta^{K,VA}_s) + va_{irst} - k_{irst}$.

Proposition 1 states an increase in M leads to a decrease in the markup for any type of firm. As such, the model entails that any first moment of the distribution of markups falls with the degree of competition. Since the median is a robust first moment, I choose $Median_{rst}[\ln \mu_{irst}]$ as the primary, inverse measure of competition at the state-sector-year level, where μ_{irst} is the plant-level markup. The measurement of μ_{irst} follows the procedure outlined by De Loecker and Warzynski (2012), and is discussed in appendix A.3. In particular, I assume plants have Cobb-Douglas production functions, minimize costs, and that labor is a variable input. Together, these assumptions imply the following expression for μ_{irst}:

$$\mu_{irst} = \beta^L_s \frac{VA_{irst}}{w_{irst} L_{irst}} \tag{2.2}$$

where $w_{irst} L_{irst}$ is the wage bill. Intuitively, when plants spend a higher share of value added on labor, conditional on the output elasticity for labor, these firms are setting a lower markup.

2.3 Stylized facts

This section first provides support for the empirical relevance of the main model assumptions, and second, presents motivating evidence in support of a central macro-level prediction of the model. To be clear, the current section does not aim to provide causal evidence. However, the next section will aim to establish a causal link between competition and plant-level capital convergence, in support of the model's predicted negative role of competition.

Validation of model assumptions

In this subsection I provide stylized facts that provide support for the empirical relevance of the assumptions that are central for generating misallocation in the two versions of the model. In one version of the model, capital misallocation arises from the interaction of financial constraints with productivity volatility. A first stylized fact will demonstrate a strong correlation between a measure for capital misallocation and measured productivity volatility. This is consistent with productivity volatility being an important driver of capital misallocation. In the second version of the model, capital misallocation arises from the birth of undercapitalized firms. The second stylized fact will document elevated capital growth for young plants, which will therefore corroborates the empirical relevance of the second version of the model.

Productivity Volatility

First, I examine the relationship between productivity volatility and dispersion in MRPK. A central mechanism in the model is that financial constraints lead to firms exhibiting delayed adjustment of their capital levels to positive productivity shocks. Asker, Collard-Wexler, and De Loecker (2014) show how in a setting with delayed adjustment of capital to productivity shocks, there is a positive relationship between the dispersion in MRPK and productivity volatility.[6] As such, documenting this positive relationship for the Indian manufacturing sector provides empirical support for the main mechanism of the model.

Asker, Collard-Wexler, and De Loecker (2014) document that such a positive relationship is significantly present across sectors within multiple countries. However, they do not analyze this relationship for the Indian ASI data, which is the dataset for this paper's empirical analysis.[7] To provide further evidence on the empirical relevance of productivity volatility and to set the stage for the capital convergence analysis in the next section, I replicate the analysis from Asker, Collard-Wexler, and De Loecker (2014) for the ASI data. Here, MRPK dispersion will be measured at the sector-year level:

$$\text{MRPK Dispersion} = Std_{st}(MRPK_{irst})$$

And the empirical measure for productivity volatility is

$$\text{Productivity Volatility} = Std_{st}(a_{it} - a_{it-1}),$$

where a_{it} is the measure of plant-level productivity. Here, a_{it} is measured as in Asker, Collard-Wexler, and De Loecker (2014), who impose that revenue takes a Cobb-Douglas form. Together with the assumption of cost-minimization these structural assumptions imply that productivity a_{it} can be measured as:

$$a_{it} = s_{irst} - \beta_s^K k_{it} - \beta_s^L l_{it} - \beta_s^M m_{it}$$

where $\beta_s^L = Median_s\left[\frac{wage\ bill_{irst}}{S_{irst}}\right]$, $\beta_s^M = Median_s\left[\frac{M_{irst}}{S_{irst}}\right]$, $\beta_s^K = 1 - \beta_s^L - \beta_s^M$. In addition, I also use a measure of productivity based on value-added:[8] $a_{irst}^{VA} = va_{irst} - \beta_s^{K,VA} k_{irst} -$

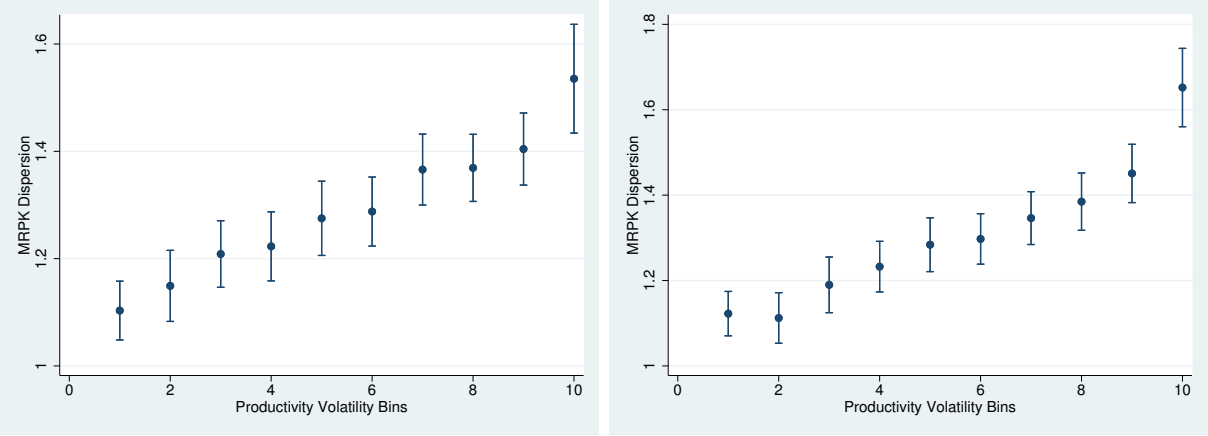
[6] Asker, Collard-Wexler, and De Loecker (2014) provide a model with capital adjustment-costs, instead of financial constraints, that also leads to delayed adjustment of capital and therefore to dispersion in MRPK. Importantly, Asker, Collard-Wexler, and De Loecker (2014) do not provide evidence for the fact that this relationship is driven by adjustment costs. Moreover, while the relationship in 2.1 is consistent with MRPK dispersion being driven by capital adjustment-costs, the evidence in the next sections, centered around the relation between capital convergence and competition, is not captured by an explanation based on adjustment costs.

[7] Asker, Collard-Wexler, and De Loecker (2014) document that this relationship holds within the Prowess dataset in India. Since Prowess features firms registered on the stock market, and consists therefore of a smaller sample than the ASI, the empirical analysis here is a useful complement to their analysis.

[8] Where value added is measured as: $VA_{irst} = S_{irst} - M_{irst}$

Figure 2.1: Productivity volatility and MRPK dispersion

(a) Gross revenue based (b) Value-added based



For the analysis in this figure, the sample is split into 10 deciles of $Std_{st}(a_{it} - a_{it-1})$. Then I run the regression $Std_{st}(MRPK_{irst}) = \sum_{D=1}^{10} \gamma_D 1(Decile\ D)_{st} + \varepsilon_{st}$, and plot the values for the coefficients and 95% confidence intervals of γ_D.

$\beta_s^{L,VA} l_{irst}$, with $\beta_s^{L,VA} = Median_s \left[\frac{wage\ bill_{irst}}{VA_{irst}}\right]$; $\beta_s^{K,VA} = 1 - \beta_s^{L,VA}$.[9] [10]

In Figure 2.1, we see that there is a strong upward sloping relationship between productivity volatility and MRPK dispersion, for both the gross-revenue based measure, and for the value-added based measure. This empirical relationship corroborates the relevance of the theoretical model.[11]

Age and Capital Growth

In an extension of the model, described in appendix A.2, I assume that firms are born with suboptimally low levels of capital. The firms' optimizing behavior then implies that, after they are born, they grow their capital to its optimal level and then remain at that capital level until they die. In this subsection, I examine whether it is empirically true in the ASI

[9] To avoid sensitivity to outliers, the median is calculated at the 2-digit sector level.

[10] Employing different productivity measures based on either gross revenue or value added serves as a primary robustness check. Since the measured elasticities for labor and capital are meaningfully different in the two measures, any sensitivity of the findings to the particular choice of output elasticities is substantially mitigated.

[11] In appendix A.4, I follow Asker, Collard-Wexler, and De Loecker (2014) by implementing variations on their plant-level robustness test for this relationship between MRPK dispersion and productivity volatility.

data that young plants exhibit higher capital growth rates than older plants.[12] [13]

In specifications 1-4 in Table 2.1, we see that the growth rate of capital is increasing with $1/age$, and therefore decreasing with age. This pattern is confirmed in specifications 5-8, as capital growth is higher for plants not older than 5 or younger than 10 years.

Table 2.1: Capital growth as a Function of Age

	Plant-level Capital Growth $g(k_{irst})$							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$\frac{1}{age_{irst}}$	0.0814**	0.140**						
	(0.00412)	(0.00569)						
$\ln(\frac{1}{age_{irst}})$			0.0136**	0.0357**				
			(0.00108)	(0.00169)				
$1(age_{irst} \leq 5)$					0.0597**	0.0576**		
					(0.00210)	(0.00281)		
$1(age_{irst} < 10)$							0.0369**	0.0306**
							(0.00190)	(0.00259)
State-sector-year FE	Yes	No	Yes	No	Yes	No	Yes	No
Plant FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	644922	644922	644922	644922	658886	658886	658886	658886

Standard errors, clustered at the plant level, in parentheses (* $p < 0.05$, ** $p < 0.01$).

Indices are i for plant, r for state, s for sector and t for year.

$g(k_{irst}) = \ln K_{irst+1} - \ln K_{irst}$, where capital is the book value of assets, measured at the start (t) and end ($t+1$) of the year.

Correlation between Competition and Misallocation

Proposition 1 states that capital wedges increase when the degree of competition is more intense. In this section, I aim to provide suggestive evidence that empirically, increased competition is associated with higher levels of measured capital misallocation. Here, the measure for capital misallocation is again the Asker, Collard-Wexler, and De Loecker (2014) measure for MRPK dispersion. The regression analysis employs the following specification:

$$Std_{rst}(MRPK_{irst}) = \gamma_s + \gamma_t + \gamma_r + \zeta Median_{rst}[\ln \mu_{irst}] + \epsilon_{rst} \tag{2.3}$$

In this specification $Median_{rst}[\ln \mu_{irst}]$ is the inverse measure of competition, $\gamma_s, \gamma_t, \gamma_r$ are sector, year and state fixed effects respectively. In alternative specifications, I also run this

[12] The existing empirical literature provides extensive support for this stylized fact, see e.g. Evans (1987), Geurts and Van Biesebroeck (2014), and Haltiwanger, Jarmin, and Miranda (2013). I here test its validity for the Indian manufacturing sector.

[13] Another relevant stylized fact relates to within-cohort capital misallocation. If there is heterogeneity across plants in capital or productivity levels at the time of their birth, translating immediately in heterogeneity in MRPK, one would expect this dispersion in MRPK to decline with age. This pattern is observed in Table **??**, and discussed in appendix A.4.

regression without γ_t, γ_r. However, I always include γ_s to eliminate variation arising from the measurement of β_s^L, the output elasticity for labor which is measured at the sector level.

Results Table 2.2 provides suggestive evidence for the prediction that MRPK dispersion might increase with competition. First we notice that $Std_{rst}(MRPK_{irst})$ is consistently negatively related to the median markup in a state-sector-year observation. This holds for both measures of $MRPK$, and it holds regardless of the specific set of fixed effects.[14]

Table 2.2: MRPK Dispersion and Competition

	$Std_{rst}(MRPK_{irst}(Gross\,Revenue))$				$Std_{rst}(MRPK_{irst}(Value\,Added))$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$Median_{rst-1}[\ln \mu_{irst-1}]$	-0.0547**	-0.0494**	-0.0371**	-0.0353**	-0.0501**	-0.0447**	-0.0376**	-0.0353**
	(0.0102)	(0.0101)	(0.0102)	(0.0101)	(0.00977)	(0.00972)	(0.0101)	(0.00997)
Constant	1.306**	1.236**	1.257**	1.228**	1.321**	1.438**	1.336**	1.477**
	(0.00487)	(0.0125)	(0.0329)	(0.0347)	(0.00465)	(0.0291)	(0.0362)	(0.0457)
Year FE	No	Yes	No	Yes	No	Yes	No	Yes
State FE	No	No	Yes	Yes	No	No	Yes	Yes
Sector FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	19951	19951	19951	19951	19570	19570	19570	19570

Standard errors, clustered at the state-sector level, in parentheses (* $p < 0.05$, ** $p < 0.01$).

Indices are i for plant, r for state, s for sector and t for year. Hence, observations are at the state-sector-year level.

Specifications 1-4 measure MRPK based on gross revenue, and specifications 5-8 based on value added.

The above evidence suggests that increased competition might worsen macro-level capital misallocation. The next empirical section will investigate the underlying micro-dynamics for the relation between competition and capital misallocation. To this end, I will examine if the model predictions on the negative link between firm-level speed of convergence and competition are observed in the data.

2.4 Competition and capital convergence

This section first explains the empirical counterparts for the model predictions, and then tests these predictions in the ASI data.

Empirical predictions

Proposition 1 summarizes the dual role for competition in the model. Since the positive role for competition is relatively standard in the literature - e.g. Peters (2013), Schaumans and Verboven (2015) - I focus here on testing the empirical presence of the negative role

[14] One might be worried about a mechanical correlation between the level of $Median_{rst-1}[\mu_{irst-1}]$ and the level of $Std_{rst}(MRPK_{irst})$. Note, however, that this would imply a positive correlation, while the regressions in Table 2.2 demonstrate a persistently negative correlation.

of competition. This negative role for competition consists in amplified capital wedges for any type of financially constrained firm in the model. These amplified capital wedges arise from slower capital growth for financially constrained firms, and therefore slower firm-level convergence to the optimal level of capital. In this section I describe the empirical tests for increased competition leading to slower capital convergence. A first set of empirical tests examines convergence in marginal revenue product of capital (MRPK), while a second set of tests examines capital growth for young plants.

Competition For both sets of empirical tests, the state-sector-year level median markup, i.e. $Median_{rst}[\ln \mu_{irst}]$, will again serve as the inverse measure of competition. Since this competition measure is arguably exogenous from the plant's point of view, this allows me to examine the causal link between competition and the empirical measures for plant-level capital convergence. Although the empirical tests will not be able to examine macro-level capital misallocation directly, an important advantage of these tests is that they can be implemented on the full panel of plants in the ASI data.

MPRK convergence In the baseline model, firms optimally choose to grow their capital stock in response to positive productivity shocks until they reach k_H^*, the optimal level of capital for high productivity firms. The empirical challenge here is that k_H^* is unobserved. To address this challenge, I focus on convergence in terms of marginal revenue product of capital (MRPK).

In terms of MRPK, the inability for a financially constrained firm to satisfy the unconstrained first-order condition in (1.14) implies that for this firm, $MRPK_{it}^* < MRPK_{it}$. Here $MRPK_{it}$ is firm i's actual MRPK in period t, and $MRPK_{it}^*$ is its optimal MRPK from the unconstrained solution. Since $MRPK_{it}$ is a strictly monotone function of k_{it}, and capital convergence in the model slows down with M, MRPK convergence also slows down with M. This is then a first empirical prediction of the model, namely that MRPK convergence is faster under lower levels of competition. In the next subsection, I will describe how I proxy for $MRPK_{it}^*$, which then allows me to analyze how convergence to $MRPK_{it}^*$ changes with the degree of competition.

Young plants In appendix section A.2, I show that a model with birth of newborn firms is isomorphic to the baseline model. As such, it has analogous implications for the rate of capital convergence as the model with productivity volatility, namely competition slows down capital convergence. The empirical advantage of this version of the model is that it allows me to test the model predictions on capital growth for young plants, which is a reduced-form object in the data.

Financial Dependence In an additional set of tests I explore the implications of heterogeneity along financial dependence for both MRPK convergence and capital growth for young plants. Here, the idea is that for sectors with higher levels of financial dependence, measured

as $Fin\ Dep_s$, changes in the level of sector-level competition have a stronger impact on the rate of MRPK convergence.

Empirically, $Fin\ Dep_s$ will be the Rajan and Zingales (1998) measure for the sector-level financial dependence. Specifically, $Fin\ Dep_s = \frac{Capital\ Expenditures_s - Cash\ Flow_s}{Capital\ Expenditures_s}$ for US sectors in the 1980's.[15] Here, $Fin\ Dep_s$ captures the share of external finance in a firm's investments in a setting with close to perfectly developed financial markets, i.e. the US. The central idea in Rajan and Zingales (1998) is then that in a setting such as India, with less developed financial markets, financial constraints become especially binding in sectors with high levels of $Fin\ Dep_s$.

Econometrics

MRPK convergence To implement the empirical test on MRPK convergence, I use the following autoregressive framework.[16]

$$MRPK_{irst} = \alpha_{irs} + \rho_0 MRPK_{irst-1} + \rho_1 MRPK_{irst-1} * Median_{rst-1}[\ln \mu_{irst-1}] + \beta X_{irst} + \gamma_t + \varepsilon_{irst} \tag{2.4}$$

The main coefficient of interest in this specification is ρ_1. This coefficient estimates how the speed of convergence changes as a function of $Median_{rst-1}[\ln \mu_{irst-1}]$.[17] To build intuition for this estimation strategy, first consider the case when $\rho_0 = \rho_1 = 0$. In that case, plants exhibit immediate convergence to the empirical proxy for $MRPK^*_{irst}$, i.e. $E[MRPK_{irst}|(\rho_0 = \rho_1 = 0)] = MRPK^*_{irst}$, regardless of $MRPK_{irst-1}$.

In practice, we will find that $0 < \rho_0 + \rho_1 < 1$ and $\rho_0 > \rho_1$, such that on average plants experience a delayed adjustment to the proxy for $MRPK^*_{irst}$. Importantly, $\rho_1 < 0$ will indicate that the speed of MRPK convergence increases with $Median_{rst-1}[\ln \mu_{irst-1}]$, as long as:

$$|\rho_0 + \rho_1 * Median_{rst-1}[\ln \mu_{irst-1}]| < \rho_0$$

I will employ two main empirical proxies for $MPRK^*_{irst}$. A first specification is $MPRK^*_{irst} = \alpha_{irs} + \beta X_{irst} + \gamma_t$, as indicated in specification 2.4, and a second measure is $MRPK^*_{irst} = \alpha_{irs} + \alpha_{rst} + \beta X_{irst}$. Here, α_{irs} is a firm-fixed effect, α_{rst} is a state-sector-year fixed effect, γ_t is a year fixed effect, and X_{irst} is a set of control variables. It is ambiguous which of the two specifications for $MRPK^*_{irst}$ is preferred, as it depends on whether year-by-year fluctuations at the state-sector level, as captured by α_{rst}, influence $MRPK^*_{irst}$ or not.[18] Throughout, the vector of control variables X_{irst} consists of a quadratic polynomial in age_{irst}.

[15] I use the ISIC Rev.2 sector definitions because these match closely with India's NIC 1987 sector definitions. The concordance between ISIC Rev.2 and NIC 1987 is provided by the Indian Statistical Office.

[16] Since $MRPK_{irst-1} * Median_{rst-1}[\ln \mu_{irst-1}]$ varies at the plant-level, standard errors will be clustered at the plant level in specifications 2.4, 2.5.

[17] In the regressions, $Median_{rst-1}[\ln \mu_{irst-1}]$ is demeaned across state-sector-year observations.

[18] To gain further understanding of the estimation procedure, note that typically a_{rst} varies over time, implying that state-sector fluctuations are correlated with $MRPK_{irst}$. The structural question is then

Financial dependence To examine the role of financial dependence in the setting of MRPK convergence, I augment the earlier specifications to allow for heterogeneous effects along financial dependence:

$$\begin{aligned} MRPK_{irst} = \alpha_{irs} + \rho_0 MRPK_{irst-1} + \rho_1 MRPK_{irst-1} * Median_{rst-1}[\ln \mu_{irst-1}] \\ + \rho_2 MRPK_{irst-1} * Fin\ Dep_s + \rho_3 MRPK_{irst-1} * Median_{rst-1}[\ln \mu_{irst-1}] * Fin\ Dep_s \\ + \beta X_{irst} + \gamma_t + \varepsilon_{irst} \end{aligned} \tag{2.5}$$

For this specification, the expectation is that $\rho_3 < 0$, as a decrease in competition would speed up convergence more for plants in sectors with higher levels of financial dependence.

Young plants The model with undercapitalized newborn plants yields empirical predictions on the capital growth for young plants. Since these predictions can be tested directly on the capital growth for young plants, a reduced-form object in the data, these tests are a useful complement to the autoregressive framework in the setting with MRPK convergence. To implement these tests, I run the following regression:[19]

$$g(k_{irst}) = \alpha_{rst} + \beta_1 young_{irst} + \beta_2 Median_{rst}[\ln(\mu_{irst-1})] * young_{irst} + \varepsilon_{irst} \tag{2.6}$$

Where I will consider three different proxies for a firm being young: $\ln(1/age_{irst}), 1(age_{irst} \leq 5), 1(age_{irst} < 10)$.

I will also examine the analogue of specification (2.5), to examine the heterogeneous effect of $Median_{rst}[\ln(\mu_{irst-1})]$ for young firms' capital growth in sectors with higher levels of financial dependence[20]:

$$\begin{aligned} g(k_{irst}) =& \alpha_{rst} + \beta_1 young_{irst} + \beta_2 Median_{rst}[\ln(\mu_{irst-1})] * young_{irst} \\ &+ \beta_3 Median_{rst}[\ln(\mu_{irst-1})] * young_{irst} * Fin\ Dep_s + \varepsilon_{irst} \end{aligned} \tag{2.9}$$

The next subsection discusses the estimation results for the above specifications.

to which extent these fluctuations influence $MRPK^*_{irst+1}$, i.e. to which extent α_{rst} is a component of $MRPK^*_{irst+1}$. Since the answer to that question is theoretically ambiguous, I perform estimations both with and without α_{rst} in the specification.

[19] For specifications 2.6, 2.9 standard errors will be clustered at the sector level.

[20] In addition to specifications (2.6), (2.9), I will also estimate:

$$g(k)_{irst} = \alpha_{irs} + \gamma_t + \beta_1 young_{irst} + \beta_2 Median_{rst}[\ln(\mu_{irst-1})] + \beta_3 Median_{rst}[\ln(\mu_{irst-1})] * young_{irst} + \varepsilon_{irst} \tag{2.7}$$

$$\begin{aligned} g(k_{irst}) =& \alpha_{irs} + \gamma_t + \beta_1 young_{irst} + \beta_2 Median_{rst}[\ln(\mu_{irst-1})] + \beta_3 Median_{rst}[\ln(\mu_{irst-1})] * young_{irst} \\ &+ \beta_4 Median_{rst}[\ln(\mu_{irst-1})] * Fin\ Dep_s + \beta_5 Median_{rst}[\ln(\mu_{irst-1})] * young_{irst} * Fin\ Dep_s + \varepsilon_{irst} \end{aligned} \tag{2.8}$$

For these two specifications, standard errors will be clustered at the plant-level.

Results

MRPK Convergence Table 2.3 provides the estimation results for specifications (2.4) and (2.5), and these results confirm the theoretical predictions of the model. First, across all specifications, in Table 2.3, the estimate for ρ_0 is both significantly different from 0 and significantly different from 1. This is consistent with the theory, which predicts that there is convergence to $MRPK^*_{irst}$ ($|\rho_0| < 1$) but that this convergence is not immediate ($\rho \neq 0$) due to financial constraints. Also note that across all specifications, the coefficient estimate for ρ_0 is on the low side. More specifically, the half-life of a deviation from $MRPK^*_{irst}$ is generally lower than 1 year. Note that this fast convergence rate lends empirical support to the choice of the proxy for $MRPK^*_{irst}$.

The focus of this empirical section is on testing the model's prediction on how competition affects the speed of convergence. For the estimated specifications, the speed of convergence always increases with $Median_{rst}[\ln(\mu_{irst-1})]$. Columns (1,2,5,6) provide the results on the baseline specifications and show that the coefficient on ρ_1 is always negative and strongly statistically significant ($p < 0.01$). This confirms the qualitative prediction of the model that the speed of convergence slows down with competition. To understand the magnitude of the estimates, examine the difference in convergence speed going from a state-sector-year observation whose median markup is in the 10th percentile of median markups, to an observation with median markup in the 90th percentile. For the baseline specification, this magnitude is largest in the specifications with both plant and state-sector-year fixed effects. For instance, in specification (2), the described comparison entails a reduction in $\rho_0 + \rho_1 * Median_{rst}[\ln(\mu_{irst-1})]$ of 0.0639, which is 19.5% of the point estimate of ρ_0.

Columns (3,4,7,8) show the results for the tests exploring heterogeneity along sectoral financial dependence. As expected, the coefficient ρ_3, estimated on the triple interaction term, is always negative. Moreover, this coefficient estimate is strongly statistically significant in columns (3,4,7).[21] [22] The estimation result that $\rho_3 < 0$ implies that the magnitude of the influence of the median markup is highest in sectors with higher financial dependence. Consider for instance a sector with a level of financial dependence at the 90th percentile in column (4). For plants in such a sector, going from a state-sector-year observation whose median markup is in the 10th percentile of median markups, to an observation with median markup in the 90th percentile of median markups, reduces $\rho_0 + ((\rho_1 + \rho_3) * Median_{rst}[\ln(\mu_{irst-1})])$ by 0.0889.

Robustness As a robustness check, appendix A.6 provides further evidence on the speed of convergence as a function of competition by analyzing convergence of the capital-labor ratio. In that appendix section, the data again confirms the predictions of the model.

[21] The exception is column (8), where the magnitude of the point estimate is comparable to those in the other specifications, although the estimate is not statistically significant.

[22] Note that the coefficient on $MRPK_{irst-1} * FinDep_s$ is always positive, which is consistent with MRPK convergence being slower in more financially dependent sectors because of the stronger salience of financial constraints.

Young Plants Table 2.4 displays the estimation results for specifications (2.6), (2.7), (2.8), (2.9). In general, capital growth for young firms increases with $Median_{rst}[\ln(\mu_{irst-1})]$, across all three measures for a firm being young, and this result is strongly statistically significant.[23] The magnitude of the point estimates is substantial. Consider again the counterfactual of moving from a state-sector-year observation whose median markup is in the 10th percentile of median markups, to an observation with median markup in the 90th percentile. For this counterfactual, the average capital growth rate increases by 3.6 percentage points for a firm less than 5 years old (specification 1).

Columns 7-12 analyze the heterogeneous effect of competition as a function of the degree of financial dependence. Across all specifications, the estimates are consistent with the theory. The heterogeneous effect is not generally statistically significant, but it is significant in columns 7 and 10 which focus on firms with $age_{irst} \leq 5$. For these specifications, the counterfactual of changing the median markup from the 10th to the 90th percentile within a sector that is at the 90th percentile of financial dependence has the substantial impact of more than 7 percentage points. This suggests that the interaction of competition with financial dependence is particularly salient for firms less than 5 years old, while still being potentially salient for slightly older firms.

Conclusion The conclusion from Tables 2.3 and 2.4 is that the data confirm that competition slows down capital convergence, both for general MRPK convergence, and for capital growth for young firms. Moreover, competition appears especially salient for capital convergence in sectors with higher levels of financial dependence.

[23]The only exception is specification (2), where the coefficient on $Median_{rst}[\ln \mu_{irst-1}] * 1(age < 10)$ is borderline significant at p= 0.084.

Table 2.3: Speed of MRPK convergence

	$MRPK_{irst}$ (Gross Revenue (GR))				$MRPK_{irst}$ (Value added (VA))			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
$MRPK_{irst-1}$ (GR)	0.442**	0.327**	0.413**	0.305**				
	(0.00571)	(0.00563)	(0.00639)	(0.00772)				
$MRPK_{irst-1}$ (GR) * $Median_{rst-1}[\ln \mu_{irst-1}]$	-0.00340**	-0.0339**	-0.000861	0.00155				
	(0.00131)	(0.0139)	(0.00213)	(0.0139)				
$MRPK_{irst-1}$ (GR) * $Fin\ Dep_s$			0.00887	0.0382*				
			(0.00572)	(0.0154)				
$MRPK_{irst-1}$ (GR) * $Median_{rst-1}[\ln \mu_{irst-1}]$ * $Fin\ Dep_s$			-0.0124**	-0.0634*				
			(0.00439)	(0.0271)				
$MRPK_{irst-1}$ (VA)					0.296**	0.189**	0.279**	0.160**
					(0.00527)	(0.00508)	(0.00613)	(0.00717)
$MRPK_{irst-1}$ (VA) * $Median_{rst-1}[\ln \mu_{irst-1}]$					-0.0103**	-0.0306**	-0.00522	0.0000582
					(0.00195)	(0.0105)	(0.00306)	(0.0144)
$MRPK_{irst-1}$ (VA) * $Fin\ Dep_s$							0.0198**	0.0593**
							(0.00688)	(0.0139)
$MRPK_{irst-1}$ (VA) * $Median_{rst-1}[\ln \mu_{irst-1}]$ * $Fin\ Dep_s$							-0.0160**	-0.0229
							(0.00612)	(0.0252)
Influence of $Median_{rst-1}[\ln \mu_{irst-1}]$ on convergence speed:								
$\rho_1 * [90\%ile[Median(\ln \mu)] - 10\%ile[Median(\ln \mu)]]$	-0.00641	-0.0639	-0.00162	0.00292	-0.0194	-0.0577	-0.00984	0.000109
$[\rho_1 + \rho_3 * Fin\ Dep_s(90\%ile)] *$ $(90\%ile[Median(\ln \mu)] - 10\%ile[Median(\ln \mu)])$			-0.0196	-0.0889			-0.0330	-0.0330
Plant FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	No	Yes	No	Yes	No	Yes	No
State-sector-year FE	No	Yes	No	Yes	No	Yes	No	Yes
Observations	238359	238359	196323	196323	196162	196162	162265	162265

Standard errors, clustered at the plant-level, in parentheses (* $p < 0.05$, ** $p < 0.01$). Variables $MRPK_{irst}$ (GR), $MRPK_{irst}$ (VA) for MRPK (marginal revenue product of capital) are defined in section 3.1. The inverse measure for competition, $Median_{rst}[\ln \mu_{irst}]$, is demeaned within sectors. All specifications include a quadratic polynomial of firm age as control variables. $90\%ile[Median(\ln \mu)]$ and $10\%ile[Median(\ln \mu)]$ are the respective values for the 90th and the 10th percentile of $Median_{rst-1}[\ln \mu_{irst-1}]$ across state-sector-year observations. This way, $\rho_1 * [90\%ile[Median(\ln \mu)] - 10\%ile[Median(\ln \mu)]]$ reports the difference in average convergence rate for firms exposed to the value of the median markup in the respective percentiles. In specifications (3,4,6,7), this is for firms in sectors with 0% financial dependence. $90\%ile[Median(\ln \mu)] * [\rho_1 + \rho_3 * Fin\ Dep_s(90\%ile)]$ $-10\%ile[Median(\ln \mu)] * [\rho_1 + \rho_3 * Fin\ Dep_s(90\%ile)]$ reports the difference in average converge rates, due to different median markups, for firms producing in sectors at the 90th percentile of financial dependence

Table 2.4: Speed of Convergence for Young Plants

	Plant-level Capital Growth $g(k_{irst})$											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$Median_{rst}[\ln \mu_{irst-1}] * 1(age \leq 5)$	0.0191**			0.0234**			0.0110			0.0118		
	(0.00649)			(0.00547)			(0.0105)			(0.00806)		
$Median_{rst}[\ln \mu_{irst-1}] * 1(age < 10)$		0.00937			0.0130*			0.00775			0.00316	
		(0.00542)			(0.00511)			(0.00858)			(0.00735)	
$Median_{rst}[\ln \mu_{irst-1}] * [[\ln(\frac{1}{age_{irst}})]$			0.00704*			0.00768**			0.00371			0.00544
			(0.00285)			(0.00257)			(0.00445)			(0.00371)
$Median_{rst}[\ln \mu_{irst-1}] * 1(age \leq 5) * Fin\ Dep_s$							0.0396*			0.0346*		
							(0.0200)			(0.0173)		
$Median_{rst}[\ln \mu_{irst-1}] * 1(age < 10) * Fin\ Dep_s$								0.0197			0.0189	
								(0.0170)			(0.0153)	
$Median_{rst}[\ln \mu_{irst-1}] * [\ln(\frac{1}{age_{irst}})] * Fin\ Dep_s$									0.0187*			0.00775
									(0.00811)			(0.00707)
$1(age \leq 5)$	0.0631**			0.0343**			0.0628**			0.0297**		
	(0.00333)			(0.00322)			(0.00358)			(0.00346)		
$1(age < 10)$		0.0385**			-0.00423			0.0395**			-0.00625	
		(0.00271)			(0.00316)			(0.00292)			(0.00343)	
$\ln(\frac{1}{age_{irst}})$			0.0142**			0.0244**			0.0151**			0.0237**
			(0.00150)			(0.00242)			(0.00164)			(0.00269)
$Median_{rst}[\ln(\mu_{irst-1})]$				0.00228	0.00289	0.0261**				0.00497	0.00679	0.0225*
				(0.00309)	(0.00348)	(0.00646)				(0.00452)	(0.00499)	(0.00933)
$Median_{rst}[\ln(\mu_{irst-1})] * Fin\ Dep_s$										-0.00555	-0.00422	0.0138
										(0.00888)	(0.00926)	(0.0182)
Influence of $Median_{rst-1}[\ln \mu_{irst-1}]$ on $g(k_{irst})$												
$\beta_2 * [90\%ile[Median(\ln \mu)] - 10\%ile[Median(\ln \mu)]]$	0.036	0.0177	0.0133	0.0441	0.0245	0.0145	0.0207	0.0146	0.006	0.0223	0.00596	0.0103
$[\beta_2 + \beta_3 * Fin\ Dep_s(90\%ile)] *$ $[90\%ile[Median(\ln \mu)] - 10\%ile[Median(\ln \mu)]]$							0.0781	0.0431	0.0341	0.0723	0.0333	0.0215
State-sector-year FE	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes	No	No	No
Plant FE, Year FE	No	No	No	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Observations	554871	554871	543752	554871	554871	543752	471868	471868	462020	471868	471868	462020

Standard errors in parentheses (* $p < 0.05$, ** $p < 0.01$).Standard errors are clustered at the state-sector level for specifications 1-3,7-9 and clustered at the plant-level for 4-6, 10-12. The inverse measure for competition, $Median_{rst}[\ln \mu_{irst}]$, is demeaned within sectors. $90\%ile[Median(\ln \mu)]$ and $10\%ile[Median(\ln \mu)]$ are the respective values for the 90th and the 10th percentile of $Median_{rst-1}[\ln \mu_{irst-1}]$ across state-sector-year observations. This way, $\beta_2 * [\ln \mu(90\%ile) - \ln \mu(10\%ile)]$ reports the difference in average capital growth for firms at the same young age level, but exposed to the different value of the median markup in the respective percentiles. In specifications (7-12), this is for firms in sectors with 0% financial dependence. Then, $\ln \mu(90\%ile) * [\beta_2 + \beta_3 * Fin\ Dep_s(90\%ile)] - \ln \mu(10\%ile) * [\beta_2 + \beta_3 * Fin\ Dep_s(90\%ile)]$ reports the difference in average capital growth, due to different median markups, for firms at the same young age level and producing in sectors at the 90th percentile of financial dependence

2.5 Competition policy reform: dereservation

In the previous section, I have analyzed the relationship between capital convergence and competition for the full panel of Indian manufacturing plants. In that setting, the identifying assumption was that the state-sector level of competition is exogenous to the individual plant. While this analysis has the advantages of employing the full panel of plants, a potential limitation is that the underlying source of the variation in competition remains unexamined. To address this concern, I now exploit natural variation in competition arising from India's 1997 dereservation reform.

Description of dereservation reform

The dereservation reform consists of the staggered removal of the small-scale industry (SSI) reservation policy. This reservation policy mandated that only industrial undertakings below a certain investment ceiling (Rs. 10 million at historical cost in 1999) were allowed to produce certain product categories.[24] In 1996, before the start of dereservation, around 1000 products were reserved for SSI.

Starting in 1997, the Indian government starts with gradually removing the reservation policy. This process of dereservation peaks between 2002 and 2008. Importantly, the timing of dereservation is arguably exogenous. A first argument for this exogeneity is given by Tewari and Wilde (2014), who document that there is considerable variation in the timing of dereservation within narrow product categories. As products within these narrow product categories arguably share the same demand and supply characteristics, this limits the scope for a structural explanation of the timing of dereservation. Moreover, Tewari and Wilde (2014) show that dereservation is uncorrelated with observable pre-policy characteristics of an industry. A more detailed description of the implementation of dereservation is provided by García-Santana and Pijoan-Mas (2014) and Martin, Nataraj, and Harrison (2014) and Tewari and Wilde (2014).

Dereservation has two distinct structural effects on incumbent plants. First, the direct effect of the removal of the investment ceiling is that incumbent establishments are allowed to grow their capital stock. Second, there is the pro-competitive shock from dereservation on incumbents. The removal of the reservation policy implies that any plant is now allowed to produce the previously reserved product. As a result, there is substantial scope for entry into the production of dereserved products. In case the pro-competitive shock is the dominant effect on a certain subset of incumbents, I can utilize the dereservation reform as an exogenous increase in the degree of competition for this subset of plants.

[24] At the time of reservation, an exception was made for large industrial undertakings already producing the product. These undertakings were allowed to continue production, but with output capped at existing levels.

Empirical analysis

Data

Data on the dereservation reform has been generously provided by Ishani Tewari, and a full description of this data is available in Tewari and Wilde (2014). Since I examine the pro-competitive effect of dereservation on incumbent plants, I will restrict the sample to plants that are observed to be incumbent at least 2 years prior to dereservation. For the purpose of this exercise, I will define a plant as being dereserved in year t if that plant's main product has been deserved during that financial year.[25]

Econometric specifications

Event study In the previous subsection, I explained how dereservation can have two opposing effects on incumbents. The direct effect of the removal of the size-cap allows plants to grow their capital, whereas the pro-competitive effect reduces profitability. In order for the dereservation reform to be relevant for the analysis in this paper, the presence of the indirect pro-competitive effect is required. To examine whether this is the case, I run the following event-study on dereservation, implemented at time $t = 0$.

$$y_{irst} = \alpha_{rs} + \gamma_t + \sum_{\tau=-4}^{4} \beta_\tau 1(t = \tau) + \varepsilon_{irst} \tag{2.10}$$

where $y_{irst} = \mu_{irst}, g(k_{irst})$ and where I bin up the end-points and normalize $\beta_{-1} = 0$.

Capital convergence After checking if the dereservation indeed has a pro-competitive impact, I examine its impact on capital convergence. This analysis is structured analogously as in section 2.4. First, I examine if the dereservation reform slows down MRPK convergence. To this end, the analogue of specification (2.4) in the dereservation setting is:

$$\begin{aligned} MRPK_{irst} = \alpha_{irs} &+ \beta_1 1(Dereserved_{irst-1}) + \rho_0 MRPK_{irst-1} \\ &+ \rho_1 MRPK_{irst-1} * 1(Dereserved_{irst-1}) + \beta_2 X_{irst} + \varepsilon_{irst} \end{aligned} \tag{2.11}$$

Here, $1(Dereserved_{irst-1})$ is an indicator variable for dereservation being implemented in period $t-1$. In case dereservation leads to slower MRPK convergence due to the pro-competitive shock, then we would expect $\rho_1 > 0$.

In addition, I examine the effect of dereservation on capital growth for young plants. To this end, I implement the analogue of specification (2.6), now in the dereservation setting:

$$\begin{aligned} g(k_{irst}) = \alpha_{irs} &+ \beta_1 (Dereserved_{irst-1}) + \beta_2 young_{irst} \\ &+ \beta_3 1(Dereserved_{irst-1}) * young_{irst} + \beta_4 X_{irst} + \varepsilon_{irst} \end{aligned} \tag{2.12}$$

[25] Since the implementation of dereservation starts in 1997, I use the NIC 1998 definition of sectors in the empirical analysis of dereservation.

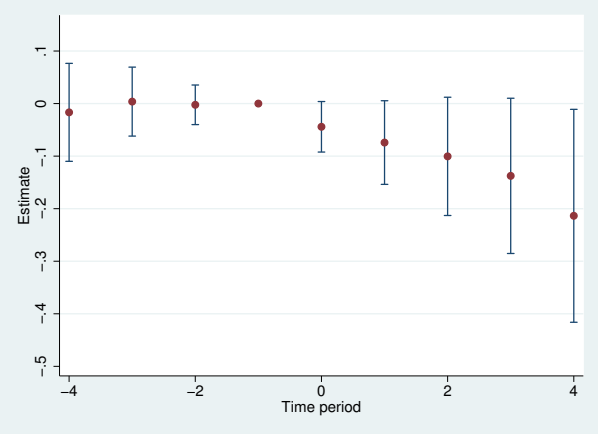
The prediction is now that $\beta_3 < 0$, in case the increase in competition due to dereservation leads to slower capital growth for young firms. The next subsection discusses the estimation results for the above specifications.

Results

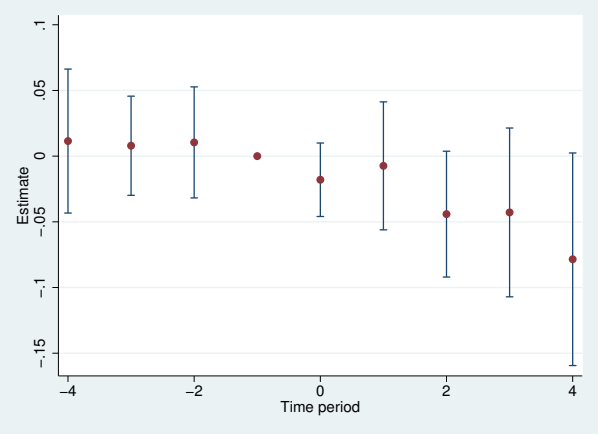
Event Study: Pro-competitive shock First, I implement specification (2.10) to examine the pro-competitive impact of dereservation on incumbent plants and Figure 2.2 shows the results.

Figure 2.2: Dereservation Event-study on Markups and Capital Growth

(a) Plant-level Markup



(b) Capital Growth



The figure displays the coefficients and 95% confidence intervals of an event-study regression on dereservation. Panels (a) displays the results of the regression $\mu_{irst} = \alpha_{irs} + \gamma_t + \sum_{\tau=-4}^{4} \beta_\tau 1(t = \tau) + \varepsilon_{irst}$, while panel (b) displays the results from the following regression: $g(k_{irst}) = \alpha_{rs} + \gamma_t + \sum_{\tau=-4}^{4} \beta_\tau 1(t = \tau) + \varepsilon_{irst}$. I impose the normalization that $\beta_{-1} = 0$.

Panel (a) indicates that plant-level markups fall after dereservation, which is consistent with a substantial pro-competitive effect of dereservation. In addition, panel (b) displays that capital growth of incumbent plants tends to fall after dereservation. However, the estimated effects are only borderline statistically significant. A possible explanation for this finding is that the direct effect of dereservation, namely the removal of the investment ceiling, partly offsets the impact of the pro-competitive shock.

In appendix A.5, I further examine if there is heterogeneity in the pro-competitive impact of dereservation. There, I find that for urban plants, which account for 62% of deserved incumbents, markups are generally lower. More importantly, I also find that dereservation leads to a more significant reduction in both markups and capital growth for urban plants compared to rural plants. Given these findings, one would expect a stronger impact of

dereservation on capital convergence for urban plants.[26] Below I examine if this expectation is confirmed by the data.

Capital convergence Table 2.5 presents the impact of dereservation on MRPK convergence. For the baseline specifications, displayed in columns (1,2,4,5), the evidence is mixed. For the gross-revenue based measure, the estimated coefficients are small and insignificant. However, the effect of dereservation on MRPK convergence in column (4) is substantial and strongly significant. Next, columns (3,6) indicate that dereservation especially slows down MRPK convergence for urban plants. This would be consistent with the findings from appendix A.5, which show that the pro-competitive impact of dereservation is particularly pronounced for urban incumbents.

Table 2.6 demonstrates that dereservation has a negative impact on the capital growth for young firms. This finding is persistent across all measures for a firm being young. Therefore, the findings on both MRPK convergence and on capital growth for young firms are consistent with the prediction, along the lines of the model, that a pro-competitive shock slows down the rate of capital convergence.

[26]Note that aside from the birth year of the plant, which is central in the analysis of capital growth for young firms, geographic location is the only other unchangeable characteristics of a plant in the data. As such, the number of degrees of freedom in the analysis of heterogeneous treatment effects is inherently limited.

Table 2.5: Speed of MRPK Convergence after Dereservation

	$MRPK_{irst}$ (Gross Revenue)			$MRPK_{irst}$ (Value added)		
	(1)	(2)	(3)	(4)	(5)	(6)
$1(Dereserved_{irst-1})$	0.0132	-0.0396	-0.0381	0.127*	0.0679	0.0607
	(0.0409)	(0.0467)	(0.0485)	(0.0611)	(0.0628)	(0.0651)
$1(Dereserved_{irst-1}) * 1(urban_{irs})$			0.0310			0.0221
			(0.0259)			(0.0311)
$MRPK_{irst-1}(GR)$	0.806**	0.497**	0.519**			
	(0.00873)	(0.0157)	(0.0160)			
$MRPK_{irst-1}(GR) * 1(Dereserved_{irst-1})$	0.00627	-0.0130	-0.0226			
	(0.0137)	(0.0151)	(0.0164)			
$MRPK_{irst-1}(GR) * 1(urban_{irs})$			-0.00173			
			(0.00642)			
$MRPK_{irst-1}(GR) * 1(Dereserved_{irst-1}) * 1(urban_{irs})$			0.0254*			
			(0.0105)			
$MRPK_{irst-1}(VA)$				0.679**	0.347**	0.350**
				(0.0111)	(0.0142)	(0.0142)
$MRPK_{irst-1}(VA) * 1(Dereserved_{irst-1})$				0.0353*	0.0213	0.00950
				(0.0154)	(0.0154)	(0.0164)
$MRPK_{irst-1}(VA) * 1(urban_{irs})$						-0.00674
						(0.00594)
$MRPK_{irst-1}(VA) * 1(Dereserved_{irst-1}) * 1(urban_{irs})$						0.0215*
						(0.00906)
State-sector FE	Yes	No	No	Yes	No	No
Plant FE	No	Yes	Yes	No	Yes	Yes
Observations	24858	25435	24294	23106	23617	23617

Standard errors in parentheses (* $p < 0.05$, ** $p < 0.01$).

All specifications include year fixed effects. Standard errors are clustered at the plant-level.

Sample includes all firms who were observed to be incumbent at least 2 years before dereservation.

Table 2.6: Speed of Convergence for Young Plants after Dereservation

	Plant-level Capital Growth $g(k_{irst})$					
	(1)	(2)	(3)	(4)	(5)	(6)
$1(Dereserved_{irst})$	-0.00916	0.0218*	-0.00620	0.0269**	-0.101**	-0.0779**
	(0.0167)	(0.00871)	(0.0176)	(0.00853)	(0.0261)	(0.0235)
$1(Dereserved_{irst}) * 1(age_{irst} \leq 5)$	-0.0705**	-0.121**				
	(0.0209)	(0.0212)				
$1(Dereserved_{irst}) * 1(age_{irst} < 10)$			-0.0429**	-0.0741**		
			(0.0123)	(0.0121)		
$1(Dereserved_{irst}) * [\ln(\frac{1}{age_{irst}})]$					-0.0316**	-0.0270**
					(0.00881)	(0.00787)
$1(age_{irst} \leq 5)$	0.0683**	0.0576**				
	(0.0158)	(0.0121)				
$1(age_{irst} < 10)$			0.0389**	0.0365**		
			(0.0117)	(0.00834)		
$\ln(\frac{1}{age_{irst}})$					0.0204*	0.0277**
					(0.00788)	(0.00625)
State-sector-year FE	Yes	No	Yes	No	Yes	No
Firm FE, year FE	No	Yes	No	Yes	No	Yes
Observations	43548	43548	43548	43548	43210	43210

Standard errors, clustered at the sector level, in parentheses (* $p < 0.05$, ** $p < 0.01$).

Sample includes all firms who were observed to be an incumbent in the reserved sector more than 2 years before dereservation.

Caveat To be clear, the evidence here should not be taken as arguing that dereservation is welfare reducing. This section is only providing evidence that dereservation has negative effects on capital convergence for incumbents, in line with the model's prediction that higher competition leads to slower convergence. A discussion of the broader (welfare) effects of dereservation can be found in García-Santana and Pijoan-Mas (2014), Martin, Nataraj, and Harrison (2014), and Tewari and Wilde (2014).

2.6 Concluding Remarks

This chapter examines the empirical implications of the theoretical model from the first chapter using Indian plant-level panel data, and it focuses on testing the negative impact of competition on capital convergence. Empirically, the prediction that the firm-level speed of capital convergence falls with competition is confirmed for the full sample of Indian manufacturing firms. This effect is particularly pronounced in sectors with higher levels of financial dependence. I also exploit natural variation in the level of competition, arising from India's 1997 dereservation reform, and again confirm the qualitative predictions of the model.

Chapter 3

Slicing the Pie: Quantifying the Aggregate and Distributional Effects of Trade

with Andrés Rodríguez-Clare and Moises Yi

3.1 Introduction

Existing gravity models of international trade provide a transparent approach to quantify the aggregate welfare effects of trade (Arkolakis, Costinot, and Rodríguez-Clare, 2012; Costinot and Rodríguez-Clare, 2014), but they remain silent on the associated distributional effects.[1] Yet, a growing empirical literature shows that trade has sharply different effects on real incomes across different groups of agents (e.g. Autor, Dorn, and Hanson (2013), Autor et al. (2014), Dix-Carneiro and Kovak (2014), and Faber (2014)). Implicitly, these two strands of the literature are reconciled by assuming that the winners compensate the losers, but then all we can say is that everybody gains from trade and not how large the social gains are. Ultimately, we want to know how the aggregate gains from trade compare with its distributional implications.

In this paper we present an integrated framework to quantify the effect of trade on the size of the pie and on the way it is sliced and divided across different groups of workers. Assuming the existence of a social welfare function, we can then further quantify the effect of trade on social welfare by adjusting for its effect on between-group inequality. The distributional effects in our model arise from a Roy (1951) structure of the labor market, where trade

[1] Notable exceptions are Fajgelbaum and Khandelwal (2014), which studies the differential effect of trade on rich and poor households, and Burstein and Vogel (2012), which analyzes the effect of trade on the skill premium.

differentially affects incomes of workers with skills that align with exportable or import-competing sectors. At the heart of the analysis is a simple expression for the change in real income due to a foreign shock (i.e. a change in trade costs or foreign technology levels) for group g in country i ,

$$\hat{W}_{ig} = \underbrace{\prod_s \hat{\lambda}_{iis}^{-\beta_{is}/\theta}}_{Multi-sector\ ACR} \cdot \underbrace{\prod_s \hat{\pi}_{igs}^{-\beta_{is}/\kappa}}_{Group-level\ Roy}, \tag{3.1}$$

where we use "hat change notation" $\hat{x} \equiv x'/x$. The first term on the right-hand side captures the change in prices given wages and is standard in the literature. As in Arkolakis, Costinot, and Rodríguez-Clare (2012) - henceforth ACR, this is given by a geometric average of the changes in the sector-level domestic trade shares elevated to the negative of the inverse of the trade elasticity, $\hat{\lambda}_{iis}^{-1/\theta}$. The second term captures the effect on the real income of group g caused by the movement in sector-level wages. It is given by a geometric average of changes in sectoral employment shares elevated to the negative of the inverse of the labor-supply elasticity to each sector, $\hat{\pi}_{igs}^{-1/\kappa}$. In our Roy model, the elasticity of labor supply to each sector, κ, is equal to the shape parameter of the Fréchet distribution that we assume governs the productivity levels that each worker draws for each sector. For both the first and second terms in Equation (3.1), the averaging weights are the Cobb-Douglas expenditure shares β_{is}.

This framework extends the existing analysis of Ricardian sector-level comparative advantage in Costinot, Donaldson, and Komunjer (2012) - henceforth CDK - to incorporate an upward sloping labor-supply curve to each sector.[2] In fact, as $\kappa \to \infty$, our model collapses to CDK. With a finite κ workers are heterogeneous in their sector-level productivities, so trade shocks that lead to the expansion of some sectors and the contraction of others have effects that vary across workers.[3] The intuition here is similar to the one in the specific-factors model. In fact, as $\kappa \to 1$ our model is equivalent to one in which workers are perfectly immobile across sectors.[4] The fact that our model nests CDK and the specific-factors model as κ moves from infinity to one implies that κ is a key parameter in the determination of the welfare effects of trade. Indeed, as we can see from Equation (3.1), given changes in sectoral

[2]CDK extend the seminal Eaton and Kortum (2002) framework to a multi-sector environment. As shown in ACR, a multi-sector version of the Armington model would be a workable substitute for the CDK-side of the model. The Krugman (1980) model or the Melitz (2003) model with a Pareto distribution (as in Chaney (2008)) would also work, though these models would introduce extra terms because of entry effects.

[3]This paper belongs to the Ricardian revival in international trade, nicely surveyed by Costinot and Vogel (2014). Their terminology of Ricardo-Roy models succinctly summarizes the framework of our model: Ricardo on the trade-side and Roy on the labor-side, capturing the source of comparative advantage at the country and worker-level respectively.

[4]For the specific-factors model (i.e., the model in which labor is sector specific), the formula in Equation (3.1) is valid for $\kappa = 1$ if we define π_{igs} as the share of earnings of group g that comes from sector s . In the Roy-Fréchet model, thinking of π_{igs} as employment shares or earning shares is equivalent. This implies that the equivalence between our model with $\kappa \to 1$ and the specific-factors model does not extend to the number of workers across sectors – in particular, for $\kappa \to 1$ the elasticity of labor supply to any particular sector with respect to the wage in that sector goes to 1 in our model but is zero in the specific-factors model.

employment shares, $\hat{\pi}_{igs}$, a lower κ implies a higher between-group variance in the welfare effects of trade shocks. The case $\kappa \to 1$ is noteworthy because then the group-level change in welfare is equal to the aggregate welfare effect multiplied by the inverse of the change in a Bartik-style index of group-level import competition.

The term labeled "Group-level Roy" in Equation [eq:gry1-1] is equal to the change in the degree of specialization of each group elevated to the power $1/\kappa$, $\hat{S}_{ig}^{1/\kappa}$, with the group-level degree of specialization S_{ig} defined as the exponential of the Kullback-Leibler divergence of the employment shares (π_{igs}, $s = 1, ..., S$) from the expenditure shares ($\beta_{is}, s = 1, ..., S$).[5] Thus, shocks that reduce a group's specialization have less beneficial welfare effects. As an example, the removal of import quotas on apparel imports from China would likely reduce the degree of specialization for a US group that specializes in apparel, exerting downward pressure on the group's welfare. Moreover, since the United States is a net importer of apparel, this group would gain from an increase in specialization if the US were to move to autarky. This formalizes the idea that groups that are specialized in import-competing sectors gain less from trade.

We use the concept of "inequality-adjusted" welfare in Jones and Klenow (2016) to measure the aggregate welfare effect of a shock that has heterogeneous effects across groups when there is no compensation for losers. One interpretation of this measure is that it captures the utility of a risk-averse agent who is behind the veil of ignorance regarding the group to which she belongs. Loosely speaking, if a shock increases inequality then the inequality-adjusted welfare effect is less favorable than the one implied by the standard aggregation, which corresponds to our measure when the coefficient of inequality aversion goes to zero.

While our methodology can be applied to several different categorizations of workers into "groups" (e.g., education, age or gender), our empirical application uses a geographical categorization. This is motivated by a growing body of empirical work documenting substantial variation in local labor-market outcomes in response to national-level trade shocks (Autor, Dorn, and Hanson, 2013; Dauth, Findeisen, and Suedekum, 2014; Dix-Carneiro and Kovak, 2014; Kovak, 2013; McLaren and Hakobyan, 2010; Topalova, 2010). Our model provides a tractable general-equilibrium framework to analyze this heterogeneous impact of trade shocks, which makes our paper a structural complement to the existing set of empirical papers.[6] We use administrative data to obtain sectoral employment shares across 15 manufacturing sectors for each of 265 regions (our groups in this application) at the Kreise level in Germany, and we combine this with data on bilateral trade flows and sectoral output from OECD STAN or the World Input-Output Database. We use this data to perform counterfactual analysis using the approach proposed by Dekle, Eaton, and Kortum (2008).

[5] Formally, $S_{ig} \equiv \exp D_{KL}(\boldsymbol{\pi}_{ig} \parallel \boldsymbol{\beta}_i) \equiv \exp \sum_s \beta_{is} \ln(\beta_{is}/\pi_{igs})$.

[6] Kovak (2013) proposes a small-economy model to understand, up to a first-order approximation, the differential effect of tariff changes across regions. Compared to that, ours is a general equilibrium model for the world economy that connects to the gravity literature and yields tractable expressions for aggregate and group-level welfare effects in terms of changes in trade and employment shares, which in turn can be computed for counterfactual shocks using the techniques in Dekle, Eaton, and Kortum (2008).

for different values of our two key parameters, θ and κ.[7]

Our first exercise is to compute the gains from trade for each region and for the country as a whole (with the standard aggregation as the population-weighted mean of regional gains), as well as the inequality-adjusted gains from trade.[8] As expected, the aggregate gains from trade and their dispersion are higher for low values of κ , with some regions actually losing from trade. Interestingly, we find that the Bartik-style index of region-level import competition perfectly predicts the ranking across regions in the gains from trade. We also find that the inequality-adjusted gains from trade are higher than the aggregate gains, as income levels become less dispersed with trade than in autarky. This is a reflection of a positive cross-region correlation in the data between earnings per worker and import competition (in manufacturing). We also find this to be the case for the United States when we use commuting zones as the definition of regions. These results suggest that trade is pro-poor in these two countries, at least from a regional perspective.

Our second exercise is to compute the welfare effects for Germany of a sector-neutral increase in productivity in China. Of particular note here is that the inequality-adjusted welfare gain is lower than the aggregate gain, a consequence of the fact that inequality across regions increases with the China shock. Hence, while trade is found to be pro-poor when compared to autarky, the rise of China is pro-rich in our simulations.

Our paper is related to several research areas in trade. In addition to the above-mentioned research on trade and local-labor markets, there is a large theoretical and empirical literature on the unequal effects of trade on labor-market outcomes – see for example Autor et al. (2014), Costinot and Vogel (2010), Burstein and Vogel (2012), Helpman et al. (2012), and Krishna, Poole, and Senses (2012). A literature focusing specifically on the effect of trade shocks on the reallocation of workers across sectors finds significant effects for developed countries (Artuç, Chaudhuri, and McLaren, 2010; Revenga, 1992),[9] which is the focus of our analysis.[10]

Artuç, Chaudhuri, and McLaren (2010) and Dix-Carneiro (2014) use a Roy model of the allocation of workers across sectors to offer a structural analysis of the dynamic adjustment to trade liberalization in a small economy. We complement these papers by linking the Roy model for the labor market with a gravity model of trade and by using the resulting framework to provide a simple and transparent way to quantify the aggregate and distributional

[7] In future work we plan to allow θ to vary across sectors. We can also allow κ to vary across groups of workers, but doing so would require estimating κ separately for each group, which is a significant challenge. We have developed extensions of our methodology to allow for intermediate goods, non-tradables or home production, and mobility of workers across regions, but at the time of writing we have not implemented these extensions in the data.

[8] As in ACR, the gains from trade are computed as the negative of the proportional welfare change caused by the country moving to autarky.

[9] See also Gourinchas (1999) and Kline (2008) for evidence of substantial reallocation in response to sectoral (but not trade) price shocks.

[10] In contrast, the evidence for developing countries suggests that reallocation in response to trade shocks is at best very sluggish – see Goldberg et al. (2007), Menezes-Filho and Muendler (2011)), and Dix-Carneiro (2014).

welfare effects of trade. Other structural analyses of trade liberalization and labor market adjustments are Coşar (2013), Coşar, Guner, and Tybout (2013), Kambourov (2009) and Ritter (2012).[11] While all these papers focus on the differential impact of trade through the earnings channel, another set of papers focuses on the expenditure channel, as in Atkin and Donaldson (2014), Faber (2014), Fajgelbaum and Khandelwal (2014) and Porto (2006).

Our paper also relates to the renewed attention to Roy models in various fields of economics – see for example Lagakos and Waugh (2013) for a recent application to development, and Young (2014) and Hsieh et al. (2013) for the productivity literature. Closer to our paper, Burstein, Morales, and Vogel (2015) utilize a Roy model with a Frchet distribution of worker abilities across occupations to decompose the changes in between-group earnings inequality into various channels, focusing on the role of technological change in explaining the evolution of the skill premium.

Finally, it is worth commenting on how our model relates to the one in Autor, Dorn, and Hanson (2013). They present a multi-sector gravity model of trade with homogeneous and perfectly mobile workers across sectors (as in CDK), but with each local economy (our groups) modeled as a separate economy. In this case all the variation in the effects of a shock across regions arise because of different terms of trade effects. In our model technologies are national and there are no trade costs among groups within countries, so terms of trade are the same for all groups. Instead, heterogeneity of workers implies that some groups of workers are more closely attached to some sectors, and it is this that generates variation in the effect of trade shocks across groups.

The rest of this paper is structured as follows. Section 3.2 provides the baseline model and its extensions. The data is described in Section 3.3, and Section 3.4 discusses our empirical findings. These empirical results include an analysis of the impact of trade-shocks on sectoral reallocation, a analysis - based on the structure of our model - of the distributional implications of a trade shock as in Autor, Dorn, and Hanson (2013), and an estimation of κ. Then, Section 3.5 presents our counterfactual analysis of a German return to autarky and of a Chinese technology shock for different values of κ. The concluding section is yet to be written.

3.2 Theory

We present a multi-sector, multi-country, Ricardian model of trade with heterogeneous workers. There are N countries and S sectors. Each sector is modeled as in Eaton and Kortum (2002 – henceforth EK): there is a continuum of goods, preferences across goods within a sector are CES with elasticity of substitution σ , and technologies have constant returns to

[11] There is also a broad literature on the impact of trade on poverty and the income distribution using a Computable General Equilibrium (CGE) methodology. Savard (2003) offers an overview of the different approaches for counterfactual analysis of the income distribution within this CGE literature, while Cockburn, Decaluwé, and Robichaud (2008) integrate multiple chapters on methodology and empirical findings of the CGE approach into a book-length discussion.

scale with productivities that are distributed Frchet with shape parameter $\theta > \sigma - 1$ and level parameters T_{is} in country i and sector s . Preferences across sectors are Cobb-Douglas with shares β_{is} . There are iceberg trade costs $\tau_{ijs} \geq 1$ to export goods in sector s from country i to country j .

On the labor side, we assume that there are G groups of workers. A worker from group g in country i has a number of efficiency units z in sector s drawn from a Frchet distribution with shape parameter $\kappa > 1$ and level parameters A_{igs} that can vary with g .[12] Thus, workers within each group are ex-ante identical but ex-post heterogeneous due to different ability draws across sectors, as in Roy (1951), while workers across groups also differ in that they draw their abilities from different distributions. The number of workers in a group is fixed and denoted by L_{ig} . In the baseline model labor supply is inelastic – workers simply choose the sector to which they supply their entire labor endowment.

If $\kappa \to \infty$ and $A_{igs} = 1$ for all g and s, the model collapses to the multi-sector EK model developed in CDK. On the other hand, if $\tau_{ijs} \to \infty$ for all j and s then economy i is in autarky and collapses to the Roy model in Lagakos and Waugh (2013) (see also Hsieh et al. (2013)). [13]

Equilibrium

To determine the equilibrium of the model, it is useful to separate the analysis into two parts: the determination of labor demand in each sector in each country as a function of wages, which comes from the EK part of the model; and the determination of labor supply to each sector in each country as a function of wages, which comes from the Roy part of the model.

Since workers are heterogeneous in their sector productivities, the supply of labor to each sector is upward sloping, and hence wages can differ across sectors. However, since technologies are national, wages cannot differ across groups. Let wages per efficiency unit in sector s of country i be denoted by w_{is}. From EK we know that the demand for efficiency units in sector s in country i is

$$\frac{1}{w_{is}} \sum_j \lambda_{ijs} \beta_{js} Y_j,$$

where Y_j is the total income for country j and λ_{ijs} are sectoral trade shares given by

$$\lambda_{ijs} = \frac{T_{is} \left(\tau_{ijs} w_{is}\right)^{-\theta}}{\sum_l T_{ls} \left(\tau_{ljs} w_{ls}\right)^{-\theta}}. \tag{3.2}$$

[12] We can easily extend the analysis to allow the Frchet parameters θ and κ to differ across sectors and groups, respectively, but choose not to do so for now to avoid notational clutter.

[13] There are two sources of comparative advantage in this model: first, as in CDK, differences in T_{is} drive sector-level (Ricardian) comparative advantage; second, differences in L/L_i and A_{igs} lead to factor-endowment driven comparative advantage. Given the nature of our comparative statics exercise, however, the source of comparative advantage will not matter for the results – only the actual sector-level specialization as revealed by the trade data will be relevant.

For future purposes, also note that the price index in sector s in country i is

$$P_{js} = \gamma^{-1} \left(\sum_i T_{is} \left(\tau_{ijs} w_{is} \right)^{-\theta} \right)^{-1/\theta}, \tag{3.3}$$

where $\gamma \equiv \Gamma(1 - \frac{\sigma-1}{\theta})^{1/(1-\sigma)}$.

Labor supply is determined by workers' choices regarding which sector to work in. Let $\boldsymbol{z} = (z_1, z_2, ..., z_S)$ and let $\Omega_s \equiv \{\boldsymbol{z} \text{ s.t. } w_{is} z_s \geq w_{ik} z_k \text{ for all } k\}$.

A worker with productivity vector $\boldsymbol{z}$ in country i will choose sector s iff $\boldsymbol{z} \in \Omega_s$. Let $F_{ig}(\boldsymbol{z})$ be the joint probability distribution of $\boldsymbol{z}$ for workers of group g in country i. The following lemma (which replicates results in Lagakos and Waugh (2013)) characterizes the labor supply side of the economy:

Lemma 3. *The share of workers in group g in country i that choose to work in sector s is*

$$\pi_{igs} \equiv \int_{\Omega_s} dF_{ig}(\boldsymbol{z}) = \frac{A_{igs} w_{is}^{\kappa}}{\Phi_{ig}^{\kappa}},$$

where $\Phi_{ig}^{\kappa} \equiv \sum_k A_{igk} w_{ik}^{\kappa}$. The supply of efficiency units by this group to sector s is given by

$$E_{igs} \equiv L_{ig} \int_{\Omega_s} z_s dF_{ig}(\boldsymbol{z}) = \frac{\eta \Phi_{ig}}{w_{is}} \pi_{igs} L_{ig},$$

where $\eta \equiv \Gamma(1 - 1/\kappa)$. [14]

One implication of this lemma is that income levels per worker are equalized across sectors. That is, for group g, we have

$$\frac{w_{is} E_{igs}}{\pi_{igs} L_{ig}} = \eta \Phi_{ig}.$$

This is a special implication of the Fréchet distribution and it implies that the share of income obtained by workers of group g in country i in sector s (i.e., $w_{is} E_{igs} / \sum w_{ik} E_{igk}$) is also given by π_{igs}. Note also that total income of group g in country i is $Y_{ig} \equiv \sum_s w_{is} E_{igs} = \eta L_{ig} \Phi_{ig}$. In turn, total income in country i is $Y_i \equiv \sum_g Y_{ig}$.

Putting the supply and demand sides of the economy together, we see that excess demand for efficiency units in sector s of country i is

$$ELD_{is} \equiv \frac{1}{w_{is}} \sum_j \lambda_{ijs} \beta_{js} Y_j - \sum_g E_{igs}. \tag{3.4}$$

Since that λ_{ijs}, Y_j and E_{igs} are functions of the whole matrix of wages $\boldsymbol{w} \equiv \{w_{is}\}$, the system $ELD_{is} = 0$ for all i, s is a system of equations in $\boldsymbol{w}$ whose solution gives the equilibrium wages for some choice of numeraire.

[14] Lemma 3 generalizes easily to a setting with correlation in workers' ability draws across sectors. In this case, the dispersion parameter κ is replaced by $\kappa/(1-\rho)$, where ρ measures the correlation parameter of ability draws across sectors for each worker. All our results below extend to this case with κ replaced $\kappa/(1-\rho)$.

Comparative Statics

Consider some change in trade costs or technology parameters. We proceed as in Dekle, Eaton, and Kortum (2008) and solve for the proportional change in the endogenous variables. Formally, using notation $\hat{x} \equiv x'/x$, we consider shocks $\hat{\tau}_{ijs}$ and $\hat{T}_{js}$ for $i \neq j$ while keeping all other parameters constant (i.e., $\hat{A}_{igs} = 1$ for all i, g, s and $\hat{L}_{ig} = 1$ for all i, g). The counterfactual equilibrium entails $ELD'_{is} = 0$ for all i, s. Noting that $w'_{is}E'_{igs} = \hat{\pi}_{igs}\hat{\Phi}_{ig}\pi_{igs}Y_{ig}$, equation $ELD'_{is} = 0$ can be written as

$$\sum_g \hat{\pi}_{igs}\hat{\Phi}_{ig}\pi_{igs}Y_{ig} = \sum_j \hat{\lambda}_{ijs}\lambda_{ijs}\beta_{js}\sum_g \hat{\Phi}_{jg}Y_{jg} \tag{3.5}$$

with

$$\hat{\Phi}_{ig} = \left(\sum_k \pi_{igk}\hat{w}_{ik}^{\kappa}\right)^{1/\kappa}, \tag{3.6}$$

$$\hat{\lambda}_{ijs} = \frac{\hat{T}_{is}\left(\hat{\tau}_{ijs}\hat{w}_{is}\right)^{-\theta}}{\sum_k \lambda_{kjs}\hat{T}_{ks}\left(\hat{\tau}_{kjs}\hat{w}_{ks}\right)^{-\theta}}, \tag{3.7}$$

and

$$\hat{\pi}_{igs} = \frac{\hat{w}_{is}^{\kappa}}{\sum_k \pi_{igk}\hat{w}_{ik}^{\kappa}}. \tag{3.8}$$

This equation can be solved for $\hat{w}_{is}$ given data on income levels, Y_{ig}, trade shares, λ_{ijs}, expenditure shares, β_{is}, labor allocation shares π_{igs}, and labor endowments, L_{ig}, and the trade-cost shocks, $\hat{\tau}_{ijs}$. From the $\hat{w}_{is}$, we can then solve for all other relevant changes, including changes in trade shares using (3.7) and changes in employment shares using (3.8).

Welfare Effects

Our measure of welfare is ex-ante real income, $W_{ig} \equiv \frac{Y_{ig}/L_{ig}}{P_i}$. We are interested in the change in W_{ig} caused by a shock to trade costs or foreign technology levels, henceforth simply referred to as a "foreign shock." Cobb-Douglas preferences combined with $Y_{ig} = \gamma L_{ig}\Phi_{ig}$ imply that

$$\hat{W}_{ig}/\hat{P}_{ig} = \hat{\Phi}_{ig}\prod_s \hat{P}_{is}^{-\beta_{is}}. \tag{3.9}$$

From (3.3) and (3.7) and given $\hat{T}_{is} = 1$ for all s we have $\hat{P}_{is} = \hat{w}_{is}\hat{\lambda}_{iis}^{1/\theta}$ while from (3.6) and (3.8) we have $\hat{w}_{is}/\hat{\Phi}_{ig} = \pi_{igs}^{1/\kappa}$. Combining these two results with (3.9) we arrive at the following proposition:

Proposition 2. *Given some shock to trade costs or foreign technology levels, the ex-ante percentage change in the real wage of group g in country i is given by*

$$\hat{W}_{ig} = \prod_s \hat{\lambda}_{iis}^{-\beta_{is}/\theta} \cdot \prod_s \hat{\pi}_{igs}^{-\beta_{is}/\kappa}. \tag{3.10}$$

The RHS of the expression in (3.10) has two components: the term $\prod_s \hat{\lambda}_{iis}^{-\beta_{is}/\theta}$ is common across groups, while all the variation across groups comes from the second term, $\prod_s \hat{\pi}_{igs}^{-\beta_{is}/\kappa}$. If $\kappa \longrightarrow \infty$, this second term converges to one, and the gains for all groups are equal to $\prod_s \hat{\lambda}_{iis}^{-\beta_{is}/\theta}$, which is the multi-sector formula for the welfare effect of a trade shock in ACR once we note that θ is the trade elasticity in all sectors in this model. It is easy to show that the term $\prod_s \hat{\lambda}_{iis}^{-\beta_{is}/\theta}$ corresponds to the change in real income given wages while the term $\prod_s \hat{\pi}_{igs}^{-\beta_{is}/\kappa}$ corresponds to the change in real income for group g coming exclusively from changes in wages $\hat{w}_{is}$ for $s = 1, ..., S$.

An alternative way to derive the result in Proposition 2 is to start from the trade and labor supply elasticities implied by our model and proceed as in ACR to infer changes in prices from trade shares and changes in wages from labor shares. Using notation $p_{ijs} \equiv w_{is}\tau_{ijs}$, the trade side of the model implies

$$\frac{d\ln(\lambda_{ijs}/\lambda_{jjs})}{d\ln(p_{ijs}/p_{jjs})} = -\theta, \tag{3.11}$$

while on the labor side we have

$$\frac{d\ln(\pi_{jgs}/\pi_{jgk})}{d\ln(w_{js}/w_{jk})} = -\kappa. \tag{3.12}$$

Envelope conditions for the consumption and work choices of agents imply

$$d\ln P_{js} = \sum_i \lambda_{ijs} d\ln p_{ijs}$$

and

$$d\ln Y_{jg} = \sum_s \pi_{jgs} d\ln w_{ijs},$$

respectively. Using $d\ln p_{jjs} = w_{js}$, solving for $d\ln p_{ijs}$ from (3.11), and plugging into the expression for $d\ln P_{js}$ yields $d\ln P_{js} = d\ln w_{js} + (1/\theta)\, d\ln \lambda_{jjs}$. Similarly, we can get $d\ln Y_{jg} = d\ln w_{js} - (1/\kappa)\, d\ln \pi_{jgs}$ for any s. Integrating these expressions yields $\hat{P}_{js} = \hat{w}_{js}\hat{\lambda}_{jjs}^{1/\theta}$ and $\hat{Y}_{jg} = \hat{w}_{js}\hat{\pi}_{jgs}^{-1/\kappa}$ for any s, and hence $\hat{Y}_{jg}/\hat{P}_{js} = \hat{\lambda}_{jjs}^{-1/\theta}\hat{\pi}_{jgs}^{-1/\kappa}$. Cobb-Douglas preferences with expenditure shares β_{js} then lead to the expression in (3.10).

The aggregate welfare effect can be obtained from Proposition 2 as $\hat{W}_i \equiv \hat{Y}_i/\hat{P}_i = \sum_g (Y_{ig}/Y_i)\hat{W}_{ig}$, where Y_{ig}/Y_i is group g's share of income. This can be written explicitly as

$$\hat{W}_i = \prod_s \hat{\lambda}_{iis}^{-\beta_{is}/\theta} \cdot \sum_g \left(\frac{Y_{ig}}{Y_i}\right) \prod_s \hat{\pi}_{igs}^{-\beta_{is}/\kappa}.$$

The aggregate welfare effect of a trade shock is no longer given by the multi-sector ACR term (i.e., $\hat{W}_i \neq \prod_s \hat{\lambda}_{iis}^{-\beta_{is}/\theta}$). This is because a trade shock will in general affect wages w_{is}, and this in turn will affect welfare through its impact on income and sector-level prices. Of course, the group level welfare effect can be seen as the product of the aggregate welfare effect and the group's relative income effect, $\hat{W}_{ig} = \hat{W}_i \cdot \left(\hat{Y}_{ig}/\hat{Y}_i\right)$. This implies

$$\frac{\hat{Y}_{ig}}{\hat{Y}_i} = \frac{\prod_s \hat{\pi}_{igs}^{-\beta_{is}/\kappa}}{\sum_h \left(\frac{Y_{ih}}{Y_i}\right) \prod_s \hat{\pi}_{ihs}^{-\beta_{is}/\kappa}}. \tag{3.13}$$

The term $\prod_s \hat{\pi}_{igs}^{-\beta_{is}/\kappa}$ is related to the change in the degree of specialization of group g. We use the Kullback-Leibler (KL) divergence as a way to define the degree of specialization of a group. Formally, the KL divergence of $\boldsymbol{\pi}_{ig} \equiv \{\pi_{ig1}, \pi_{ig2}, ..., \pi_{igS}\}$ from $\boldsymbol{\beta}_i \equiv \{\beta_{i1}, \beta_{i2}, ..., \beta_{iS}\}$ is given by

$$D_{KL}(\boldsymbol{\pi}_{ig} \parallel \boldsymbol{\beta}_i) \equiv \sum_s \beta_{is} \ln(\beta_{is}/\pi_{igs}).$$

Note that if group g in country i was in full autarky (i.e., not trading with any other group or country) then $\pi_{igs} = \beta_{is}$. Thus, $D_{KL}(\boldsymbol{\pi}_{ig} \parallel \boldsymbol{\beta}_i)$ is a measure of the degree of specialization as reflected in the divergence of the actual distribution $\boldsymbol{\pi}_{ig}$ relative to $\boldsymbol{\beta}_i$. We can now write

$$\prod_s \hat{\pi}_{igs}^{-\beta_{is}/\kappa} = \exp\left(\frac{1}{\kappa}\left[D_{KL}(\boldsymbol{\pi}'_{ig} \parallel \boldsymbol{\beta}_i) - D_{KL}(\boldsymbol{\pi}_{ig} \parallel \boldsymbol{\beta}_i)\right]\right).$$

This implies that the welfare effect of a trade shock on a particular group is determined by the change in the degree of specialization of that group as measured by the KL divergence (modulo $\prod_s \hat{\lambda}_{iis}^{-\beta_{is}/\theta}$). Consider a group g in country i that happens to have efficiency parameters $(A_{ig1}, ..., A_{igS})$ that give it a strong comparative advantage in a sector s for which the country as a whole has a comparative disadvantage, as reflected in positive net imports in that sector. Group g would be highly specialized in s when the country is in autarky (but groups trade among themselves) but that specialization would diminish as the country starts trading with the rest of the world. As a consequence, the KL degree of specialization falls with trade for group g, implying lower gains relative to other groups in the economy.

Gains from Trade

Following ACR, we define the gains from trade as the negative of the proportional change in real income for a shock that takes the economy back to autarky: $GT_i \equiv 1 - \hat{W}_i^A$ and

$GT_{gi} \equiv 1 - \hat{W}_{ig}^{A}$. A move to autarky for country i entails $\hat{\tau}_{ijs} = \infty$ for all s and all $i \neq j$. Conveniently, solving for changes in wages in country i (i.e., solving for $\hat{w}_{is}$ for $s = 1, ..., S$) from Equation (3.5) only requires knowing the values of trade and employment shares for country i, namely λ_{iis} for all s and π_{igs} for all g, s. This can be seen by letting $\hat{\tau}_{ijs} \to \infty$ in Equation (3.5), which yields

$$\sum_{g} \hat{\pi}_{igs} \hat{\Phi}_{ig} \pi_{igs} Y_{ig} = \beta_{is} \sum_{g} \hat{\Phi}_{ig} Y_{ig}. \tag{3.14}$$

Proposition 3. *For a finite κ, the aggregate gains from trade are higher than those in the model with $\kappa \longrightarrow \infty$.*

To understand this result, it is useful to consider the simpler case with a single group of workers, $G = 1$. For a move back to autarky, in this case we would have

$$\hat{W}_{i}^{A} = \prod_{s} \lambda_{iis}^{\beta_{is}/\theta} \cdot \exp\left[-\frac{1}{\kappa} D_{KL}(\boldsymbol{\pi}_i \parallel \boldsymbol{\beta}_i) \right].$$

Since $D_{KL}(\boldsymbol{\pi}_i \parallel \boldsymbol{\beta}_i) > 0$, then (given $\boldsymbol{\pi}_i$) a lower κ implies a lower $\hat{W}_i$. Intuitively, a finite κ introduces more "curvature" to the PPF, making it harder for the economy to adjust as it moves to autarky. This implies higher losses if the economy were to move to autarky, and hence higher gains from trade, – see Costinot and Rodríguez-Clare (2014). Proposition 2 establishes that this result generalizes to the case $G > 1$.

Turning to the group-specific gains from trade, we again use the KL measure of specialization to understand whether a group gains more or less than the economy as a whole. The results of the previous section imply that the gains from trade for group g in country i are

$$GT_{ig} = 1 - \prod_{s} \lambda_{iis}^{\beta_{is}/\theta} \cdot \exp\left(\frac{1}{\kappa} \left[D_{KL}(\boldsymbol{\pi}_{ig}^{A} \parallel \boldsymbol{\beta}_i) - D_{KL}(\boldsymbol{\pi}_{ig} \parallel \boldsymbol{\beta}_i) \right] \right).$$

The term $D_{KL}(\boldsymbol{\pi}_{ig}^{A} \parallel \boldsymbol{\beta}_i) - D_{KL}(\boldsymbol{\pi}_{ig} \parallel \boldsymbol{\beta}_i)$ could be positive or negative, depending on whether group g in country i becomes more or less specialized with trade as measured by the KL divergence. Intuitively, if a group happens to be specialized in industries that face strong import competition, this would imply that $D_{KL}(\boldsymbol{\pi}_{ig} \parallel \boldsymbol{\beta}_i) < D_{KL}(\boldsymbol{\pi}_{ig}^{A} \parallel \boldsymbol{\beta}_i)$, and hence lower gains from trade.

A Limit Case

An interesting case arises in the limit as $\kappa \to 1$, where the model becomes isomorphic to one in which labor cannot move across sectors (i.e., where L_{igs} is fixed). In this case we can easily get that for a foreign shock we have

$$\lim_{\kappa \to 1} \hat{Y}_{ig} = \sum_{s} \pi_{igs} \hat{w}_{is}. \tag{3.15}$$

Letting $r_{is} \equiv \sum_g \pi_{igs} Y_{ig}/Y_i$ be the share of sector s in total output in country i, we then have

$$\lim_{\kappa \to 1} \hat{r}_{is} = \frac{\hat{w}_{is}}{\sum_k r_{ik} \hat{w}_{ik}}.$$

Combined with $\lim_{\kappa \to 1} \hat{Y}_i = \sum_k r_{ik} \hat{w}_{ik}$ we finally get

$$\lim_{\kappa \to 1} \hat{Y}_{ig}/\hat{Y}_i = \sum_s \pi_{igs} \hat{r}_{is}. \tag{3.16}$$

The benefit of this result is that $\hat{r}_{is}$ is observable in the data. Thus, if we can identify the impact of a foreign shock on output shares, then we can compute the implied relative income changes across groups. As in Autor, Dorn, and Hanson (2013), we can instrument the Bartik-style variable $\sum_s \pi_{igs} \hat{r}_{is}$ with $\sum_s \pi_{igs} \Delta IP_s^{East \to Other}$ and run an IV regression of observed $\hat{Y}_{ig}/\hat{Y}_i$ on $\sum_s \pi_{igs} \hat{r}_{is}$. If κ was indeed very close to 1 then this regression should yield a coefficient close to one. We expect that the coefficient will be lower than one precisely because workers can and would move across sectors in response to relative wage changes.[15] We will check this result in the empirical section.

The case $\kappa \to 1$ also leads to a sharp result for the change in relative income levels across groups in a move back to autarky. From Equation (3.14) combined with (3.8) we get

$$\hat{w}_{is}^{\kappa} \sum_g \hat{\Phi}_{ig}^{1-\kappa} \pi_{igs} Y_{ig} = \beta_{is} \hat{Y}_i Y_i$$

Setting $\hat{Y}_i = 1$ by choice of numeraire and setting $\kappa = 1$ yields $\hat{w}_{is} = \beta_{is}/r_{is}$. Plugging into (3.15) yields

$$\lim_{\kappa \to 1} \frac{\hat{Y}_{ig}^A}{\hat{Y}_i^A} = I_{ig} \equiv \sum_s \pi_{igs} \frac{\beta_{is}}{r_{is}}. \tag{3.17}$$

We can think of β_{is}/r_{is} as an index of the degree of import competition in industry s and I_{ig} as an index of import competition faced by group g. Thus, in the limit as $\kappa \to 1$, the change in relative income levels across groups is simply given by the index of import competition that we can directly observe in the data.Things are more complicated in the general case with $\kappa > 1$, but we will see that I_{ig} remains a good proxy for whether $\hat{Y}_{ig}^A/\hat{Y}_i^A \gtrless 1$ and that the variance of $\hat{Y}_{ig}^A/\hat{Y}_i^A$ across g falls with κ. Of course, one can also use the result in (3.17) to rewrite the result in (3.16) and get an expression for any foreign shock as

$$\lim_{\kappa \to 1} \frac{\hat{Y}_{ig}}{\hat{Y}_i} = \frac{1}{\hat{I}_{ig}}. \tag{3.18}$$

[15]The coefficient could be higher than one if there is mobility across regions or if the labor supply to the manufacturing sector is not perfectly inelastic. Below we present extensions of the model to allow for these two possibilities, which we plan to explore quantitatively in the near future.

Inequality-Adjusted Welfare Effects

Consider an agent "behind the veil of ignorance" who doesn't know what group she will belong to. Since there are L_{ig} workers in group g, the probability that our agent behind the veil will end up in group g is $l_{ig} \equiv L_{ig}/L_i$. Let ρ denote the degree of relative risk aversion. The certainty-equivalent real income of an agent behind the veil is

$$U_i \equiv \left(\sum_g l_{ig} W_{ig}^{1-\rho} \right)^{1/(1-\rho)}.$$

We can think of $V_i \equiv W_i/U_i$ as a measure of the cost of inequality for an agent behind the veil of ignorance. Consistent with this idea, V_i is equal to one if $\rho = 0$ and is increasing in ρ, reaching $W_i/\min_g W_{ig}$ when $\rho \to \infty$.[16]

In the quantitative section below we will present results for "inequality-adjusted" welfare effects of a foreign shock, defined as $\hat{U}_i$ for any foreign shock, and "inequality-adjusted" gains from trade, defined as $IGT_i \equiv 1 - \hat{U}_i^A$ for a shock that takes the economy back to autarky.[17]

We will compare these effects with the standard ones, $\hat{W}_i$ and $GT_i = 1 - \hat{W}_i^A$. Given our definition of V_i, we have $\hat{U}_i = \hat{W}_i/\hat{V}_i$, $IGT_i = 1 - \frac{1-GT_i}{\hat{V}_i^A}$. If the foreign shock increases inequality ($\hat{V}_i > 1$) then $\hat{U}_i < \hat{W}_i$ while if inequality falls ($\hat{V}_i < 1$) then $\hat{U}_i > \hat{W}_i$. Similarly, if inequality is higher in the observed equilibrium than in autarky then $IGT_i < GT_i$, while in the opposite case $IGT_i > GT_i$.

Alternative Models and Extensions

In this section we extend the model to allow for an upward sloping labor supply to the whole manufacturing sector (Section 3.2), intermediate goods (Section 3.2), and mobility across groups, which is particularly relevant to the case in which groups correspond to geographic regions (Section 3.2).

Upward sloping labor supply

We extend the model by introducing a new sector in which goods can only be traded within each group. This non-tradable sector is identical to all other sectors regarding the labor and technology dimensions, with the main difference being that the elasticity of substitution in consumption between this sector and the rest can be different than one. As we show next,

[16] Related welfare measures are examined by Cordoba and Verdier (2008) and Heathcote, Storesletten, and Violante (2008) and Jones and Klenow (2016), who incorporate income risk into the analysis of aggregate welfare in macro models without trade. Antras, Gortari, and Itskhoki (2016) introduce a measure of inequality-adjusted gains from trade that is closely related to ours, but their focus is on analyzing the role for redistribution after trade liberalization in a setting with distortionary income tax-transfer system.

[17] These inequality-adjusted welfare effects focus on between-group inequality. For within-group inequality, the model implies that the distribution of worker income q follows $Pr(q \leq Q) = e^{-\Phi_g^{\kappa} Q^{-\kappa}}$. Hence, inequality measures which are invariant to the scale of the Fréchet are unaffected by the trade shocks.

if the elasticity of substitution between tradables and non-tradables is higher than one then the labor supply to the tradable sector is increasing in the real wage in the tradable sector. We discuss this further below.

The wage in the non-tradable sector (indexed by $s = 0$) can differ across groups (i.e., $w_{ig0} \neq w_{ih0}$ for $g \neq h$). Lemma 3 still applies,[18] and the equilibrium system is similar to what we had above, except that now expenditure shares vary across groups. Letting ξ_{ig} denote the share of total expenditure in tradables for group g in country i, the excess labor demand in sector $s \geq 1$ is now

$$ELD_{is} \equiv \frac{1}{w_{is}} \sum_{j,g} \lambda_{ijs} \beta_{js} \xi_{jg} Y_{jg} - \sum_{g} E_{igs},$$

while in sector $s = 0$ the excess labor demand in group g in country i is

$$ELD_{ig0} = \frac{1}{w_{ig0}} (1 - \xi_{ig}) Y_{ig} - E_{igs}.$$

In turn, letting χ denote the elasticity of substitution in consumption between tradables and non-tradables, the expenditure shares on tradable goods are given by

$$\xi_{ig} = \frac{\left(\prod_{s \geq 1} P_{is}^{\beta_{is}} \right)^{1-\chi}}{\left(\prod_{s \geq 1} P_{is}^{\beta_{is}} \right)^{1-\chi} + P_{ig0}^{1-\chi}},$$

with P_{is} for $s \geq 1$ still given (3.3) and $P_{ig0} = \eta^{-1} T_{i0}^{-1/\theta} w_{ig0}$. Without loss of generality we assume henceforth that $T_{i0} = \eta^{-\theta}$ for all i, so that $P_{ig0} = w_{ig0}$.

Noting that $\lambda_{ijs}, \xi_{ig}, Y_{jg}$ and E_{igs} are all functions of the matrix of wages $\boldsymbol{w^T} \equiv \{w_{is}\}$ for all i and $s = 1, ..., S$ and the vector $w^{NT} \equiv \{w_{ig0}\}$ for all ig, the system $ELD_{is} = 0$ for all i, s and ELD_{ig0} for all ig is a system of equations in $\boldsymbol{w^T}$ and w^{NT} whose solution gives the equilibrium wages for some choice of numeraire. We can proceed as above and write down the equations for the hat changes in wages given some shock to trade costs or technology levels – see the Appendix for details. Here we are interested in showing how the value of χ determines the slope of the labor supply to the tradable sector.

The condition $ELD_{ig0} = 0$ is simply $1 - \xi_{ig} = \pi_{ig0}$. Assuming without loss of generality that $A_{i0} = 1$ for all i, and letting $w_{igM} \equiv \left(\sum_{s \geq 1} A_{igs} w_{is}^{\kappa} \right)^{1/\kappa}$, this can be rewritten as

$$\frac{\left(w_{ig0} / w_{igM} \right)^{1-\chi}}{\left(w_{igM} / \prod_{s \geq 1} P_{is}^{\beta_{is}} \right)^{\chi - 1} + \left(w_{ig0} / w_{igM} \right)^{1-\chi}} = \frac{\left(w_{ig0} / w_{igM} \right)^{\kappa}}{1 + \left(w_{ig0} / w_{igM} \right)^{\kappa}}$$

[18] Using notation w_{igs} for wages for convenience (in equilibrium we still have $w_{igs} = w_{is}$ for all $s \geq 1$), employment shares are now $\pi_{igs} = A_{igs} w_{igs}^{\kappa} / \Phi_{ig}$ while the supply of efficiency units is $E_{igs} = \frac{\gamma \Phi_{ig}}{w_{igs}} \pi_{igs} L_{ig}$, with $\Phi_{ig} \equiv \sum_s A_{igs} w_{igs}^{\kappa}$.

. If $\chi > 1$ then the LHS is decreasing in w_{ig0}/w_{igM} (demand curve) while the RHS is increasing in w_{ig0}/w_{igM} (supply curve). A decrease in the real manufacturing wage $w_{igM}/\prod_{s\geq 1} P_{is}^{\beta_{is}}$ implies shift to the right of the demand curve, leading to an increase in the equilibrium w_{ig0}/w_{igM} and an increase in π_{ig0}. Thus, a shock that decreases the real manufacturing wage also leads to an increase in the share of people that move into the non-tradable sector.

As mentioned above, we think of the addition of the non-tradable sector as a particularly convenient way to get the labor supply curve to the manufacturing sector to be upward sloping in the real manufacturing wage. This requires that $\chi > 1$, which could seem contrary to the standard custom in the international macroeconomics literature to assume that the elasticity of substitution between tradables and non-tradables is lower than one. However, the non-tradable sector also includes "home production," and a recent literature in macroeconomics argues that adjustment in hours devoted to home production may explain the variation in market hours over the business cycle, with a central value of $\chi = 2.5$ (Aguiar, Hurst, and Karabarbounis, 2013). Since we can think of our sector $s = 0$ as including both standard non-tradables as well as home production, it is reasonable to assume $\chi > 1$.

Intermediate goods

Consider again the basic model but now with an input-output structure as in Caliendo and Parro (2014). This extension is important because a significant share of the value of production in a sector originates from other sectors, and taking this into account may affect the effects of trade on wages $\hat{w}_{is}$ and hence the welfare effects across groups.

The labor supply of the model is exactly as in the main model (as characterized by Lemma 3), and trade shares and the price indices are given as in (**??**) and (3.3), except that instead of w_{is} we now have c_{is}, where c_{is} is given by

$$c_{is} = w_{is}^{1-\alpha_{is}} \prod_k P_{ik}^{\alpha_{iks}}. \tag{3.19}$$

Here the α_{iks} are the Cobb-Douglas input shares: a share α_{iks} of the output of industry s in country i is used buying inputs from industry k, and $1 - \alpha_{is}$ is the share spent on labor, with $\alpha_{is} = \sum_k \alpha_{iks}$. Combining this expression for c_{is} with (3.3) (but with w_{is} replaced by c_{is}) yields

$$P_{js} = \eta^{-1} \left(\sum_i T_{is} \tau_{ijs}^{-\theta} w_{is}^{-(1-\alpha_{is})\theta} \prod_k \left(P_{ik}^{-\theta}\right)^{\alpha_{iks}} \right)^{-1/\theta}.$$

Given wages, this equation represents a system of NxS equations in P_{js} for all j and s, which can be used to solve for P_{js} and hence c_{is} and λ_{ijs}. This implies that trade shares are an implicit function of wages.

Let X_{js} and R_{js} be total expenditure and total revenues for country j on sector s. We know that $R_{is} = \sum_{j=1}^{n} \lambda_{ijs} X_{js}$ while Cobb-Douglas preferences and technologies imply

$X_{js} = \beta_{js} Y_j + \sum_{k=1}^{S} \alpha_{jsk} R_{jk}$. Combining these equations we get a system of linear equations that we can use to solve for revenues given income levels and trade shares,

$$R_{is} = \sum_j \lambda_{ijs} \left(\beta_{js} Y_j + \sum_{k=1}^{S} \alpha_{jsk} R_{jk} \right).$$

Since trade shares and income levels themselves are a function of wages, this implies that revenues are a function of wages. The excess demand for efficiency units in sector s of country i is now

$$ELD_{is} \equiv R_{is} - \sum_g E_{igs}.$$

As in the baseline model, the system $ELD_{is} = 0$ for all i, s is a system of equations that we can use to solve for wages. In turn, given wages we can solve for all the other variables of the model.

The next step is to write the hat algebra system. From $ELD'_{is} = 0$ we get

$$\sum_g \hat{\pi}_{igs} \hat{\Phi}_{ig} \pi_{igs} Y_{ig} = (1 - \alpha_{is}) \sum_{j=1}^{n} \lambda_{ijs} \hat{\lambda}_{ijs} \left(\beta_{js} \sum_g \hat{\Phi}_{jg} Y_{jg} + \sum_{k=1}^{S} \alpha_{jsk} \hat{R}_{jk} R_{jk} \right),$$

where $\hat{\Phi}_{ig}$ is as in (3.6) and

$$\hat{\lambda}_{ijs} = \frac{\left(\hat{\tau}_{ijs} \hat{w}_{is}^{1-\alpha_{is}} \prod_k \hat{P}_{ik}^{\alpha_{iks}} \right)^{-\theta}}{\sum_l \lambda_{ljs} \left(\hat{\tau}_{ljs} \hat{w}_{ls}^{1-\alpha_{ls}} \prod_k \hat{P}_{lk}^{\alpha_{lks}} \right)^{-\theta}},$$

$$\hat{P}_{js}^{-\theta} = \sum_i \lambda_{ijs} \hat{\tau}_{ijs}^{-\theta} \hat{w}_{is}^{-(1-\alpha_{is})\theta} \prod_k \left(\hat{P}_{ik}^{-\theta} \right)^{\alpha_{iks}},$$

and

$$\hat{R}_{is} R_{is} = \sum_j \lambda_{ijs} \hat{\lambda}_{ijs} \left(\beta_{js} \sum_g \hat{\Phi}_{ig} Y_{jg} + \sum_{k=1}^{S} \alpha_{jsk} \hat{R}_{jk} R_{jk} \right).$$

Analogous to Proposition 2, from the hat algebra we find the following result:

Proposition 4. *Given some trade shock, the ex-ante percentage change in the real wage of group g in country i is given by*

$$\hat{W}_{ig} = \prod_{s,k} \hat{\lambda}_{iik}^{-\beta_{is} \widetilde{a}_{isk}/\theta} \cdot \prod_{s,k} \hat{\pi}_{igk}^{-\beta_{is} \widetilde{a}_{isk}(1-\alpha_{ik})/\kappa} \tag{3.20}$$

where $\widetilde{a}_{i,sk}$ is the typical element of matrix $\left(I - \Upsilon_i^T\right)^{-1}$ with $\Upsilon_i \equiv \{\alpha_{iks}\}_{k,s=1,\ldots,S}$.

Mobility Across Regions

In our model, the ability of workers can be interpreted as being determined by the fundamentals of the region where they work, in addition to innate characteristics particular to the worker's region of origin.[19] Under this interpretation, workers have an incentive to move across regions in response to trade shocks, which is something we have not modeled thus far.[20] Here we consider an extension of the benchmark model where workers can move across regions but not across countries. Assume that each worker gets a draw in each sector and each region. Workers also have an "origin region." We say that a worker with origin region g is "from region g." Each worker gets a draw z in each region-sector combination (h, s) from a Fréchet distribution with parameters κ and A_{ihs}. Workers are fully described by a matrix $\boldsymbol{z} = \{z_{hs}\}$ and an origin region g. A worker from region g in country i that wants to work in region h of country i suffers a proportional adjustment to income determined by ζ_{igh}, with $\zeta_{igg} = 1$ and $\zeta_{igh} \leq 1$ for all i, g, h. Thus, a worker from g who works in region h in sector s has income of $w_{is}\zeta_{igh}z_{hs}$.

We now let

$$\Omega_{igfs} \equiv \{\boldsymbol{z} \text{ s.t. } w_{is}\gamma_{igf}z_{fs} \geq w_{ik}\gamma_{igh}z_{hk} \text{ for all } h, k\}.$$

A worker with productivity matrix $\boldsymbol{z}$ from region g in country i will choose region-sector (f, s) iff $\boldsymbol{z} \in \Omega_{igfs}$. The following lemma characterizes the labor supply side of the economy:

Lemma 4. *The share of workers in group g in country i that choose to work in (f, s) is*

$$\pi_{igfs} \equiv \int_{\Omega_{gfs}} dF(\boldsymbol{z}) = \frac{A_{fs}\left(\zeta_{gf}w_{is}\right)^{\kappa}}{\Phi_{ig}^{\kappa}},$$

where $\Phi_{ig}^{\kappa} \equiv \sum_{h,k} A_{hk}\left(\zeta_{gh}w_{ik}\right)^{\kappa}$. The efficiency units supplied by this group in sector (f, s) are given by

$$E_{igfs} \equiv L_{ig}\int_{\Omega_{gfs}} z_{fs}dF_i(z) = \pi_{igfs}\gamma L_{ig}\frac{\Phi_{ig}}{w_{is}\zeta_{igf}}.$$

[19] Specifically, there are two ways to interpret our baseline model. First, one could think that the $\boldsymbol{z}$ is inherent to the worker, something that the worker is born with, and that if she were to migrate to another region this $\boldsymbol{z}$ would not change. Since wages vary across sectors but not across regions, this interpretation would imply that there are no incentives for workers to migrate. Second, one could think that all workers draw an x in each sector from a Fréchet distribution with parameters 1 and κ, and that their efficiency units if they work in (g, s) are $A_{igs}^{1/\kappa}x_s$ (note that this is isomorphic to our current specification because $\Pr(z \leq a) = \Pr(A_{igs}^{1/\kappa}x \leq a)$). In this interpretation, $A_{igs}^{1/\kappa}$ is a region-sector specific shifter that is common to all workers, and x is an worker-specific idiosyncratic term that is distributed the same everywhere. If we adopt the second interpretation, then labor income would differ across regions for the same worker, and there would be an incentive to migrate. For example, workers would want to move to regions that have a comparatively high common shifter in sectors whose relative wage increases after the trade shock.

[20] There is limited empirical evidence of geographic mobility in response to trade shocks. Autor, Dorn, and Hanson (2013), Dauth, Findeisen, and Suedekum (2014), and Topalova (2010) find that trade shocks induced only small population shifts across regions in the US, Germany, and India, respectively. These studies focus on the short and medium run, while ours focuses on the long run.

Total income of group g in country i is $Y_{ig} \equiv \sum_{f,s} w_{is}\zeta_{gf}E_{igfs} = \gamma L_{ig}\Phi_{ig}$. Moreover, the share of income obtained by workers in group g in country i in region-sector (f,s) is also given by π_{igfs}, while (ex-ante) per capita income for workers of group g in country i is $Y_{ig}/L_{ig} = \gamma\Phi_{ig}$.

Let $\mu_{igh} \equiv \sum_s \pi_{ighs}$ be the share of workers from g that work in h. It is easy to verify that $\pi_{ighs}/\mu_{igh} = \pi_{ihhs}/\mu_{ihh}$ for all i, g, h, s. Thus, conditional on locating in region h, all workers irrespective of their origin have sector employment shares given by $\pi_{ihs} \equiv \pi_{ighs}/\mu_{igh}$. The shares π_{ihs} and μ_{igh} will be enough to characterize the equilibrium below.

The labor demand side of the model is exactly as in the case with no labor mobility across regions. Putting the supply and demand sides of the economy together, we see that excess demand for efficiency units in sector s of country i is

$$ELD_{is} \equiv \frac{1}{w_{is}}\sum_j \lambda_{ijs}\beta_{js}Y_j - \sum_{g,h} E_{ighs}.$$

Noting that λ_{ijs}, Y_j and E_{ighs} are functions of the whole matrix of wages $\boldsymbol{w} \equiv \{w_{is}\}$, the system $ELD_{is} = 0$ for all i, s is a system of equations in $\boldsymbol{w}$ whose solution gives the the equilibrium wages for a given choice of numeraire. Turning to comparative statics, the implications of a trade shock can be characterized in similar fashion to what we did in Section 3.2. Changes in wages can be obtained as the solution to the system of equations given by

$$\sum_{g,h} \hat{\pi}_{ihs}\hat{\Phi}_{ig}\mu_{igh}\pi_{ihs}Y_{ig} = \sum_j \lambda_{ijs}\hat{\lambda}_{ijs}\beta_{is}\sum_g \hat{\Phi}_{jg}Y_{jg} \tag{3.21}$$

with $\hat{\Phi}_{ig}^{\kappa} = \sum_{h,s}\mu_{igh}\pi_{ihs}\hat{w}_{is}^{\kappa}$, (3.7) and $\hat{\pi}_{ihs} = \hat{\pi}_{ighs}/\hat{\mu}_{igh}$, $\hat{\pi}_{ighs} = \hat{w}_{is}^{\kappa}/\hat{\Phi}_{ig}^{\kappa}$, and $\hat{\mu}_{igh} = \sum_s \pi_{ihs}\hat{\pi}_{ighs}$. Equation (3.5) can be solved for $\hat{w}_{is}$ given data on income levels, Y_{ig}, trade shares, λ_{ijs}, migration shares μ_{igh} , employment shares π_{ihs} , and the shocks, $\hat{\tau}_{ijs}$ and $\hat{T}_{js}$. In turn, given $\hat{w}_{is}$, changes in trade shares can be obtained from (3.7), while changes in migration and employment shares can be obtained from the expressions for $\hat{\pi}_{ihs}$ and $\hat{\mu}_{igh}$ above.

Given $\hat{w}_{ik}$, the following proposition analogous to Proposition 2 characterizes the impact of a trade shock on ex-ante real wages for different groups of workers.

Proposition 5. *Given some trade shock, the ex-ante percentage change in the real wage of group g in country i is given by* $\hat{W}_{ig} = \prod_s \hat{\lambda}_{iis}^{-\beta_{is}/\theta} \cdot \prod_s \left(\hat{\mu}_{igg}\hat{\pi}_{igs}\right)^{-\beta_{is}/\kappa}$

For the limit case $\kappa \to 1$ we again have $\lim_{\kappa\to 1}\hat{Y}_{ig}/\hat{Y}_i = 1/\hat{I}_{ig}$, except that now at $I_g \equiv \sum_s v_{igs}\frac{\beta_{is}}{r_{is}}$, where $\nu_{igs} \equiv \sum_h \mu_{igh}\pi_{ihs}$ is the share of workers from region g that work in sector s.

3.3 Data

National figures on bilateral trade flows, sectoral output and employment shares come mostly from the OECD Database for Structural Analysis (STAN), and are supplemented with data from the World Input-Output Database (WIOD). For Germany, we obtain regional employment shares (π_{igs}) and output shares (needed to compute r_{is}) using data from the German Social Security System. For reasons of convenience we restrict our simulation analysis to the year 2003. We are in the process of reproducing the simulations for other years. Our choice of industry classification is driven by the availability of the data. We aggregate manufacturing industries into 15 groups which roughly correspond to two-digit ISIC Rev. 3 codes ($S = 15$).

Table 3.1: List of Industries

ISIC Rev. 3 Code	Description
15-16	C15T16 Food products, beverages and tobacco
17-19	C17T19 Textiles, textile products, leather and footwear
20	C20 Wood and products of wood and cork
21-22	C21T22 Pulp, paper, paper products, printing and publishing
23	C23 Coke, refined petroleum products and nuclear fuel
24	C24 Chemicals and chemical products
25	C25 Rubber and plastics products
26	C26 Other non-metallic mineral products
27	C27 Basic metals
28	C28 Fabricated metal products, except machinery and equipment
29	C29 Machinery and equipment, n.e.c.
30-33	C30T33 Electrical and optical equipment
34	C34 Motor vehicles, trailers and semi-trailers
35	C35 Other transport equipment
36-37	C36T37 Manufacturing n.e.c. and recycling

For Germany, the geographical units of observation g are German Kreise, which are roughly equivalent to US counties. Each of these regions contains a minimum of 100,000 inhabitants as of December of 2008. In the current version of the data, we observe 265 of these regions (all located in West Germany).[21] [22]

Our measures of trade flows are taken from the OECD-STAN database. To arrive at our measures, we combine values of national sectoral output,[23] and total import and export

[21] The employment counts are based on the job in which workers spent the longest spell during 2003.

[22] In cases where $\pi_{igs} = 0$, we imputed a small value to make the data consistent with our model.

[23] Output measures Y_{is} are based on STAN variable PROD "Production (gross output)" (see Appendix for detailed description). We acknowledge that there is a mismatch between the labor data, which corresponds to West German regions, and the trade data, which corresponds to the whole of Germany. We will work on improving this in the near future.

figures by sector. This allows us to obtain consistent values of import penetration by sector:

$$\lambda_{iis} = \frac{Y_{is} - X_{is}^{WORLD}}{Y_{is} - X_{is}^{WORLD} + M_{is}^{WORLD}}$$

Where $X_{is}^{WORLD}, M_{is}^{WORLD}$ are exports and imports in country i, sector s, to or from the rest of the world. We then obtain the consumption shares β_{is} as follows:

$$\beta_{is} = \frac{Y_{is} - X_{is}^{WORLD} + M_{is}^{WORLD}}{\sum_s Y_{is} - X_{is}^{WORLD} + M_{is}^{WORLD}}$$

In our estimations in Section 3.4, we supplement our trade figures with data from the United Nations Commodity Trade Statistics Database (UN Comtrade) in order to obtain instrumental variables for region-level import penetration consistent with the work by (Dauth, Findeisen, and Suedekum, 2014).

3.4 Empirics

In this section, we bring the model to the data with a set of empirical exercises. First, we examine if trade shocks lead to regional sectoral reallocation. Since the theory predicts that the distributional impact of trade across groups depends on how a region's comparative advantage corresponds to the pattern of trade-induced sectoral expansion and contraction, documenting that trade shocks lead to sectoral reallocation is a first step in establishing the fit of the model to the data. In a second exercise, we examine how trade-shocks affect income changes at the regional level. More specifically, we test the model-based prediction on how regional changes in import competition affect regional income levels and the findings from this exercise will validate our theory on this dimension, and at the same time establish our model as a structural framework for understanding the existing empirical findings on the distributional impact of trade across local labor markets (see e.g. Autor, Dorn, and Hanson (2013)). In a third exercise, we then estimate κ, a central structural parameter in our model. This estimation strategy will exploit the structural relationship between sectoral reallocation and income changes at the region level and thereby synthesize the first two empirical exercises.

Trade and Sectoral Reallocation

Here, we check in the data if trade shocks indeed lead to sectoral reallocation. This is mainly a sanity check to see if a central mechanism in our theoretical framework - trade-induced sectoral reallocation - is indeed present in the data. While a number of existing studies document substantial sectoral reallocation in response to trade shocks[24], it is still worthwhile to examine if this pattern also holds for our specific empirical setting. Moreover, the patterns explored in the empirical exercises in the next subsections will also relate to sectoral reallocation, and thereby the current exercise helps for understanding the setting of our broader empirical analysis.

Measuring trade shocks

As a measure of trade shocks, we employ changes in import penetration from the East. Following the intuition of Autor et al. (2014) (and its application by Dauth, Findeisen, and Suedekum (2014) to Germany), changes in sectoral trade flows from China and Eastern Europe[25] to a group of countries "similar" to Germany[26] proxy for changes in sectoral import-

[24]Relevant papers are e.g. Artuç, Chaudhuri, and McLaren (2010) and Revenga (1992), as mentioned in the introduction.

[25]Eastern Europe is comprised of the following countries: Bulgaria, Czech Republic, Hungary, Poland, Romania, Slovakia, Slovenia, and the former USSR or its succession states Russian Federation, Belarus, Estonia, Latvia, Lithuania, Moldova, Ukraine, Azerbaijan, Georgia, Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan, and Uzbekistan.

[26]We follow Dauth, Findeisen, and Suedekum (2014) in defining this set of countries to include Australia, Canada, Japan, Norway, New Zealand, Sweden, Singapore, and the United Kingdom. Countries were selected

competition from the East. The idea behind this approach is to capture the effect of changes in trade-flows driven by sector-specific growth in China and Eastern Europe, which are plausibly exogenous to changes in fundamentals on the demand or supply side in Germany.[27] Specifically, our trade-shock measure is

$$\Delta IP_{st}^{East \to Other} \equiv \frac{\Delta M_{st}^{East \to Other}}{L_{st}^{Germany}},$$

where $L_{st}^{Germany}$ is the number of workers in Germany employed in industry s at the beginning of time period t and $M_{st}^{East \to Other}$ are the net imports by the above-defined set of countries similar to Germany from China and Eastern Europe.[28]

Region-level sectoral reallocation

After defining our measure for trade shocks, we are now able to examine if trade causes region-level sectoral reallocation.[29]To this end, we run the following regression:

$$\Delta \ln \tilde{\pi}_{igst} = \gamma \Delta IP_{st}^{East \to Other} + \zeta_{st} \tag{3.22}$$

with $\tilde{\pi}_{igst} \equiv \frac{\pi_{igst}}{\pi_{igs_0t}}$.[30] Table 3.2 presents the estimation results for this regression. For each of the specifications, we find a negative impact of increased import penetration on the relative growth of a sector, and except for lags 5 and 6, this impact is significant at the 5% level. This confirms that our trade-shock variable $\Delta IP_{st}^{East \to Other}$ induces sectoral reallocation in the expected direction. In terms of magnitude, a 1000 euro increase in $\Delta IP_{st}^{East \to Other}$ leads to a decrease in $\tilde{\pi}_{igst}$ of circa 0.4 log points.

based on having a similar income level as Germany, but all direct neighbors and members of the European Monetary Union were excluded.

[27] The intuition behind the instrument is that the "rise of the East" is an exogenous event, affecting trade for all countries at comparable levels of development as Germany in a similar way. For a discussion on the exogeneity restrictions and the robustness of this type of instrument, see Dauth, Findeisen, and Suedekum (2014), as well as Autor, Dorn, and Hanson (2013).

[28]The number of years over which $\Delta M_{st}^{East \to Other}$ is computed will vary, and is defined in the regression table.

[29]We focus on region-level sectoral reallocation, as this is the level of reallocation most relevant from the model's perspective. Appendix section B.1 complements the analysis in this section by examining reallocation at the national level. There, we first decompose Germany-wide sectoral reallocation, in terms of output shares, to examine which share of the reallocation is trade-related. We find that 64% of the variance of changes in output shares is due to changes in trade-related reallocation. Since the correlation between changes in output shares and changes in employment shares is 56.8%, this is additional, strongly suggestive evidence that trade-shocks can lead to substantial sectoral reallocation.

[30]For each group g there are $S-1$ degrees of freedom for the reallocation of π_{igst}. This is why we normalize to a reference sector π_{igs_0t}. Also, since there can be no common trend across s for $\Delta \ln \tilde{\pi}_{igst}$, we do not include a constant in equation (3.22).

Table 3.2: Labor Reallocation in Response to Trade Shock

	$\Delta \ln \tilde{\pi}_{igst}$						
	Lag = 3	Lag = 4	Lag = 5	Lag = 6	Lag = 7	Lag = 8	Lag = 9
$\Delta IP_{st}^{East \to Other}$	-0.0063***	-0.0053**	-0.0038*	-0.0039*	-0.0042**	-0.0041**	-0.0042**
	[0.0090]	[0.0220]	[0.0749]	[0.0519]	[0.0120]	[0.0340]	[0.0440]
Observations	4018	4009	4002	3999	3988	3983	3987

Because of low number of clusters, we show symmetric p-values from wild cluster bootstrap-t Wald test in brackets, with 1,000 replications, clustered by 14 industry cells. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$ Data from 1999-2008; $\Delta IP_{st}^{East \to Other} = \frac{\Delta M_{st}^{East \to Other}}{L_{st}}$ with $\Delta M_{st}^{East \to Other}$ in 1000 EURO, with base year 2005. The number of lags indicates the number of years in one time period over which we compute the LHS and RHS variables. Lags 3 and 4 allow for multiple time-periods. Estimations with a higher number of lags only include the time period starting in 1999, with base year 2005. All estimations use sector 1516 as the numeraire sector for the dependent variable.

Regional changes in import competition and income

In this subsection, we analyze how trade shocks affect income changes at the regional level. More specifically, we explore the empirical relationship between regional income changes and our model's measure for changes in import competition. Based on the analysis in Section 3.2, using equation (3.16) and $\hat{W}_{ig} = \hat{W}_i \hat{Y}_{ig} / \hat{Y}_i$, we know that in the limit as $\kappa \to 1$, we have that

$$\ln \hat{W}_{ig} = \ln \hat{W}_i + \ln \sum_s \pi_{igs} \hat{r}_{is}. \tag{3.23}$$

Moreover, the simulation-based analysis in Section 3.5 will confirm that for values of $\kappa > 1$, the relationship between $\hat{W}_{ig}$ and $\sum_s \pi_{igs} \hat{r}_{is}$ remains close to log-linear. We now bring this theoretical relationship to the data with the following regression specification, where we allow the coefficient on $\ln \sum_s \pi_{igs} \hat{r}_{is}$ to be empirically determined. :

$$\ln \hat{Y}_{ig} = \alpha + \beta \ln \sum_s \pi_{igs} \hat{r}_{is} + \varepsilon_{ig} \tag{3.24}$$

Through the lens of our model, the error term ε_{ig} could be driven by a change in A_{ig}, measurement error or a deviation of the log-linear relationship in (3.23) for values of $\kappa > 1$. We instrument for $\ln \sum_s \pi_{igs} \hat{r}_{is}$ with $\sum_s \pi_{igs} \Delta IP_{st}^{East \to Other}$, which turns the above defined sector-level trade-shock into a Bartik (1991) style instrument for trade shocks at the regional level.

Table 3.3 presents the results.[31] The estimated coefficient is positive and strongly statistically significant in all specifications except one. This result corroborates the theoretical prediction that trade-induced changes in regional income depend positively on $\sum_s \pi_{ig} \hat{r}_{is}$.

[31] Appendix Figure B.10 plots the scatters for the first stage.

Table 3.3: Changes in import-competition and regional income per worker

	First Stage						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Lag = 3	Lag = 4	Lag = 5	Lag = 6	Lag = 7	Lag = 8	Lag = 9
$\sum_s \pi_{igs} \Delta IP_s^{East \to Other}$	-0.001***	-0.002***	-0.005***	-0.005***	-0.004***	-0.004***	-0.003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.51	0.77	0.47	0.44	0.38	0.33	0.24
F-stat	8.26	87.37	135.56	174.69	152.07	178.67	106.43
$\ln \hat{Y}_{ig}$	**Second Stage**						
$\sum_s \pi_{igs} \hat{r}_{is}$	1.220*	0.440**	0.206	0.267**	0.398***	0.344**	0.515***
	(0.662)	(0.181)	(0.137)	(0.114)	(0.130)	(0.136)	(0.173)
Observations	795	530	265	265	265	265	265

Standard errors, clustered at the group-level, in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$
Data from 1999-2008, with $\Delta IP_s^{East \to Other}$ denoted in 1000 EURO with base year 2005. The number of lags indicates the number of years in one time period over which we compute the LHS and RHS variables. Lags 3 and 4 allow for multiple time-periods. Estimations with a higher number of lags only include the time period starting in 1999.

Finally, the confidence intervals for the coefficients in the more precisely estimated specifications are contained within zero and one, as required by the theory (see discussion in Section 3.2.[32]) In addition to corroborating our theoretical framework, the findings in Table 3.3 are in line with the existing evidence on the regional impact of increased import competition, documented by Autor, Dorn, and Hanson (2013), Dauth, Findeisen, and Suedekum (2014), Dix-Carneiro and Kovak (2014), Kovak (2013), and McLaren and Hakobyan (2010) and Topalova (2010).

Estimation of κ

As is evident from equation (3.1), the κ parameter is central to our model as it jointly affects the aggregate and the distributional welfare-effects from trade.[33] In this subsection, we will structurally estimate κ. As we explain below, the estimation procedure will exploit the trade-induced relationship between sectoral reallocation and real-income changes, both at the region-level. From this perspective, this section also synthesizes our empirical analysis of the impact of trade shocks on regional sectoral reallocation and regional changes in real income.

Based on Burstein, Morales, and Vogel (2015), our estimation approach relies on the structural relationship between regional income changes $\hat{Y}_{ig}$ and sectoral reallocation. The model implies that unobserved changes in relative wages result in relative changes in sectoral

[32]Below, the simulations in Section 3.5 will also illustrate this requirement.

[33]We have imposed that θ, the main other structural parameter, is equal across sectors. Relaxing this assumption would affect the aggregate gains of trade, but not the distribution of gains. For discussion and estimation of θ, see Caliendo and Parro (2014) and Head and Mayer (2014)

shares, which are observable. Since regional income changes are also a function of sectoral wages, it is then intuitive that there is also a structural relation between these regional income changes and regional sectoral reallocation. Moreover, this structural relationship between these two observables is governed by κ, and this relationship can then be exploited to structurally estimate κ.

Formally, the relationship between income and employment at the region-level can be written as,

$$\hat{Y}_{ig} = \left(\sum_s \pi_{igs} \hat{w}_{is}^{\kappa} \right)^{1/\kappa}$$

We also have $\hat{\pi}_{igs} = \frac{\hat{w}_{is}^{\kappa}}{\hat{\Phi}_{ig}^{\kappa}}$ and hence

$$\frac{\hat{w}_{is}^{\kappa}}{\hat{w}_{i1}^{\kappa}} = \frac{\hat{\pi}_{igs}}{\hat{\pi}_{ig1}} \tag{3.25}$$

Combining these expressions we obtain the following equation

$$\ln \hat{Y}_{ig} = \frac{1}{\kappa} \ln \left(\sum_s \pi_{igs} \frac{\hat{\pi}_{igs}}{\hat{\pi}_{ig1}} \hat{w}_{i1}^{\kappa} \right) \tag{3.26}$$

Before we take this equation to the data[34], we reduce sensitivity to group-level noise by observing that equation (3.25) holds for all g in country i such that we can define $\nu_{is}(k) \equiv \exp \frac{1}{G} \sum_g \log \frac{\hat{\pi}_{igs}}{\hat{\pi}_{igk}}$ and then update equation (3.26) to

$$\ln \hat{Y}_{ig} = \frac{1}{\kappa} \ln \left(\sum_s \pi_{igs} \nu_{is}(1) \hat{w}_{i1}^{\kappa} \right)$$

To eliminate the sensitivity of this relation to the choice of reference sector, we use the fact that $\forall k : \hat{w}_{is}^{\kappa} = \hat{w}_{ik}^{\kappa} \nu_{is}(k)$ and write $\hat{w}_{is}^{\kappa} = \left(\exp \frac{1}{S} \sum_k \log \hat{w}_{ik}^{\kappa} \right) \nu_{is}$, where $\nu_{is} \equiv \exp \left(\frac{1}{S} \sum_k \log \nu_{is}(k) \right)$. This way, we arrive at the following equation

$$\sum_s \pi_{igs} \hat{w}_{is}^{\kappa} = \left(\exp \frac{1}{S} \sum_k \log \hat{w}_{ik}^{\kappa} \right) \sum_s \pi_{igs} \nu_{is} \tag{3.27}$$

We can then substitute equation (3.27) into (3.26) and obtain our estimating equation,

$$\ln \hat{Y}_{ig} = b_i + \frac{1}{\kappa} \ln \sum_s \pi_{igs} \nu_{is} + \varepsilon_{ig} \tag{3.28}$$

[34]Note that it simplifies to $\ln \hat{Y}_{ig} = a_i - \frac{1}{\kappa} \ln \hat{\pi}_{ig1}$,with $a_i = \ln \hat{w}_{i1}^{\kappa}$. This equation can be taken to the data, but is sensitive to the choice of reference sector.

where $b_i \equiv \frac{1}{\kappa}\left(\frac{1}{S}\sum_k \log \hat{w}_{ik}^{\kappa}\right)$. Finally, we require exogenous variation in $\sum_s \pi_{igs}\nu_{is}$. To this end, we use the Bartik-type instrument $\sum_s \pi_{igs}\Delta IP_{st}^{East\to Other}$, as explained in the previous subsection.

Table 3.4 presents the results for our κ estimation.[35] The point estimates for κ range from 2.9 to 5, with a 95% confidence interval for the most precise κ estimate (specification 5) of 1.1-4.7. For the next section, where we will run simulations to analyze the quantitative role of κ in our framework, we will set our preferred point estimate at $\kappa = 3$.

Table 3.4: Reallocation and regional income per worker

	First Stage						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Lag = 3	Lag = 4	Lag = 5	Lag = 6	Lag = 7	Lag = 8	Lag = 9
$\sum_s \pi_{igs}\Delta IP_s^{East\to Other}$	-0.004***	-0.005***	-0.005***	-0.004***	-0.005***	-0.005***	-0.005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
R^2	0.54	0.60	0.65	0.55	0.60	0.69	0.70
F-stat	263.64	178.24	232.19	125.45	145.62	219.13	300.04
$\ln \widehat{Y}_{ig}$	**Second Stage**						
$\ln \sum_s \pi_{igs}\nu_{is}$	0.299***	0.224***	0.202	0.293**	0.342***	0.249***	0.292***
	(0.086)	(0.083)	(0.134)	(0.127)	(0.108)	(0.094)	(0.085)
Implied κ	3.346***	4.463***	4.957	3.413**	2.927***	4.019***	3.419***
	(0.958)	(1.659)	(3.305)	(1.478)	(0.926)	(1.518)	(0.997)
Observations	795	530	265	265	265	265	265

Standard errors, clustered at the group-level, in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
Standard errors for the implied κ computed using the Delta method.
Data from 1999-2008, with $\Delta IP_s^{East\to Other}$ denoted in 1000 EURO with base year 2005. The lag after 1999 indicates the construction of the time-period. Lags 3 and 4 allow for multiple time-periods.
$\widehat{Y}_{ig}$ is measured as average income per manufacturing worker.

[35] Appendix Figure B.9 plots the scatters for the first stage.

3.5 Counterfactual Simulations

Using our baseline model and the methodology described in Section 3.2, in this section we perform two counterfactual exercises: a move to autarky by Germany and a sector-neutral productivity increase in China. For each of these two cases, we compute group-level, aggregate and inequality-adjusted welfare effects, $\hat{W}_{ig}, \hat{W}_i$ and $\hat{U}_i$, respectively, for $i = Germany$ and $g = 1, ..., 265$. In all the ensuing excercises, we follow Costinot and Rodríguez-Clare (2014) in assuming a value of $\theta = 5$, which is the central value for the trade elasticity as reviewed in Head and Mayer (2014).

A Move to Autarky

Table **??** summarizes the results for $\hat{W}_{ig}$ and $\hat{W}_i$. For a value of $\kappa = 3$, our results indicate an aggregate loss of 11.4%, with a significant dispersion in these losses across regions (the coefficient of variation is 0.438). The loss from a return to autarky decreases with κ, with an aggregate loss of 13.9% when $\kappa \to 1$ and 8% when $\kappa \to \infty$. The intuition is that a lower κ introduces more curvature to the PPF, making it harder to adjust to autarky. Note that for $\kappa \to 1$, some regions gain substantially from going to autarky (and therefore lose from trade). Even for $\kappa = 15$, there are still regions that gain from a return to autarky.

Figure 3.1 plots the distribution of regional losses for different values of κ. A lower κ leads to higher dispersion in these losses due to a stronger pattern of worker-level comparative advantage. As κ approaches infinity, workers are perfectly substitutable across sectors, and the variance in regional gains from trade gradually disappears.

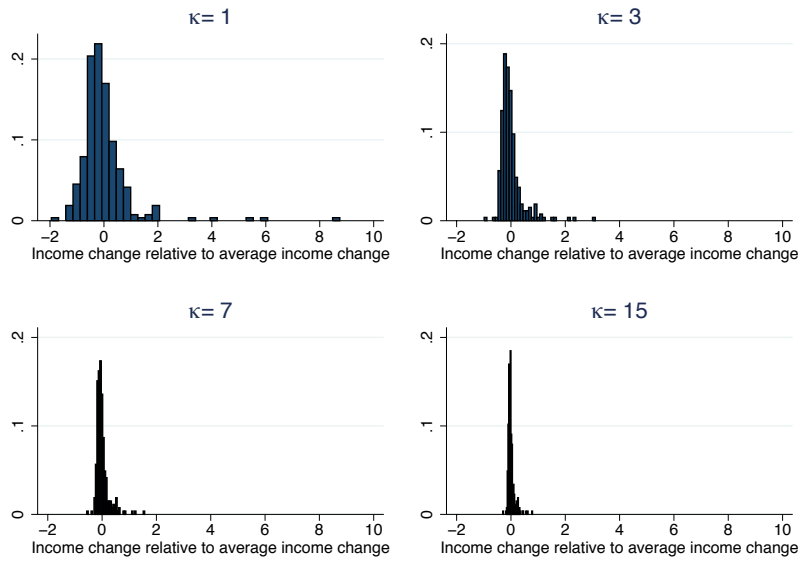
Table 3.5: Germany's return to autarky: summary statics

	$\overline{\widehat{W}_{ig}}$ (Δ %)	$\widehat{W}_i$ (Δ %)	CV	Min.	Max.
$\kappa \to 1$ (Specific Factors)	-12.4	- 13.9	0.992	-1.949	8.744
$\kappa = 3$	-11.8	-11.4	0.438	-0.988	3.105
$\kappa = 7$	-9.5	-9.9	0.235	-0.565	1.567
$\kappa = 15$	-8.7	-9.2	0.126	-0.299	0.793
$\kappa \to \infty$ (CDK)	-8.0	-8.0	0	1	1
N		265			

$\overline{\widehat{W}_{ig}} \equiv \frac{1}{G}\sum_g \widehat{W}_{ig} - 1$ and the first and second column display $100(\overline{\widehat{W}_{ig}} - 1), 100(\widehat{W}_i - 1)$ respectively. The third column displays the coefficient of variation (CV). For the final two columns: Min.= $[\min_g(\widehat{W}_{ig}) - 1]/[\overline{\widehat{W}_{ig}} - 1]$ and Max.=$[\max_g(\widehat{W}_{ig}) - 1]/[\overline{\widehat{W}_{ig}} - 1]$.

In our simulations, regions specialized in import-competing sectors tend to lose less than export-oriented regions. Employment shares in the autarky equilibrium are given by expenditure shares, $r^a_{is} = \beta_{is}$, so the ratio β_{is}/r_{is} proxies the necessary expansion/contraction that a sector has to undergo at the national level as country i moves to autarky. We can

Figure 3.1: Germany moves to autarky: distribution of gains by region



then think of this ratio as a sector-level index of import competition, with $\beta_{is}/r_{is} > 1$ (< 1) indicating an import-competing (export-oriented) sector. Table 3.6 shows that this index varies considerably across manufacturing industries in Germany, reaching a maximum for sector 23, "Coke, refined petroleum products and nuclear fuel," with $\beta_{is}/r_{is} = 9.16$, and a minimum for sector 29, "Machinery and equipment," with $\beta_{is}/r_{is} = 0.65$. Taken together, this sizable variation in β_{is}/r_{is} implies considerable sectoral reallocation under a return to autarky.

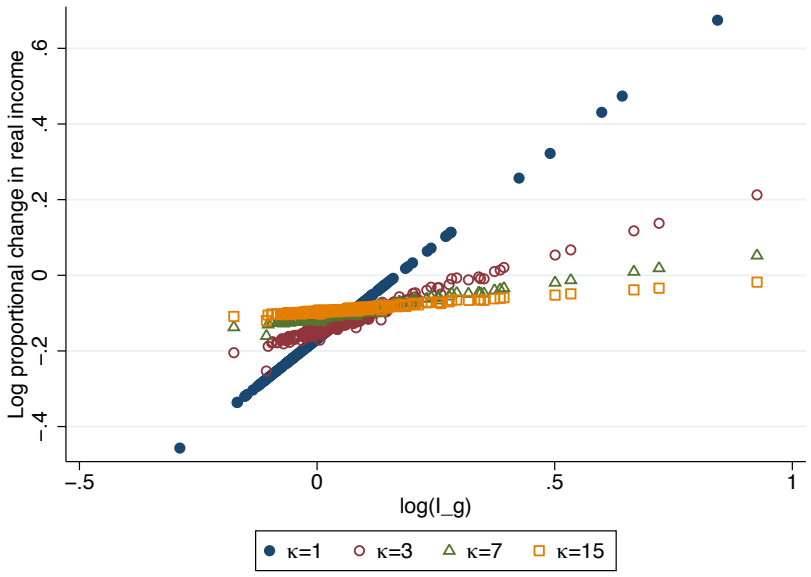
Figure 3.2 presents the results for group-level gains from trade (in logs, vertical axis) against the Bartik-style region-level index of import competition defined in Section 3.2 , $I_{ig} \equiv \sum_g \pi_{igs} \frac{\beta_{is}}{r_{is}}$ (in logs, horizontal axis). In Section 3.2 we showed that in the limit as $\kappa \to 1$ this index perfectly captures the variation in group-level gains from trade, which is confirmed by the slope of 1 in the points corresponding to this case. The figure also shows that although the slope is no longer one when $\kappa > 1$, the correlation between $\log \hat{W}_{ig}$ and $\log I_{ig}$ is almost one, indicating that the I_{ig} does a very good job in ranking regions according to their gains from trade. These simulation results thereby provide further model-based support for the regression analysis in Section 3.4, where we examined the impact of regional changes in important competition on regional income.

Naturally, the regions that lose from trade are the most import-competing regions. Here, Gelsenkirchen is the region that is most affected, with an increase in real income of 23.7% from moving to autarky when $\kappa = 3$. This is mainly because it has 18% of its manufacturing workforce employed in sector 23 "Coke, refined petroleum products and nuclear fuel."

Table 3.6: Index of sectoral import competition

β_{is}/r_{is}	**Industry**
1.224	s = C15T16 Food products, beverages and tobacco
1.26	s = C17T19 Textiles, textile products, leather and footwear
0.865	s = C20 Wood and products of wood and cork
0.838	s = C21T22 Pulp, paper, paper products, printing and publishing
9.159	s = C23 Coke, refined petroleum products and nuclear fuel
1.342	s = C24 Chemicals and chemical products
0.715	s = C25 Rubber and plastics products
0.989	s = C26 Other non-metallic mineral products
1.11	s = C27 Basic metals
0.706	s = C28 Fabricated metal products, except mach. and equip.
0.647	s = C29 Machinery and equipment, n.e.c.
0.93	s = C30T33 Electrical and optical equipment
1.408	s = C34 Motor vehicles, trailers and semi-trailers
1.162	s = C35 Other transport equipment
0.826	s = C36T37 Manufacturing n.e.c. and recycling

Figure 3.2: Distribution of Gains by Region



Given the distribution of group-level gains from trade, we can compute the inequality-adjusted gains from trade (IGT), as described in Section 3.2. Figure 3.3 shows that for a strictly positive coefficient of relative risk aversion, the IGT for Germany are higher than the standard aggregate gains from trade. Loosely speaking, this comes from the fact that there is less inequality across regions with trade than in the autarky counterfactual. For $\kappa = 3$, the gains from trade are 11.2% while $IGT = 12.8\%$ for a coefficient of inequality aversion of 2. Furthermore, the IGT tends to increase, though not monotonically, with the coefficient of inequality aversion.

In Figure 3.4 we provide some insight into why $IGT > GT$. In the data, the correlation between import-competition and average earnings per worker is positive at 0.33, which explains why trade is on average pro-poor. In addition, the bottom percentiles of the income distribution pre-dominantly feature export-oriented regions, and these regions gain more from trade than the average regionThis means that certainly for high ρ, $IGT > GT$.

For the US, we find broadly similar patters, with IGT larger than regular gains of trade, and increasingly so for higher coefficients of relative risk aversion. These patterns are displayed in Figure B.7.

Productivity increase in China

Baseline results

Motivated by recent research on the rise of China and its distributional impact on US (Autor, Dorn, and Hanson, 2013) or German (Dauth, Findeisen, and Suedekum, 2014) labor markets,

Figure 3.3: Inequality-adjusted Gains from Trade

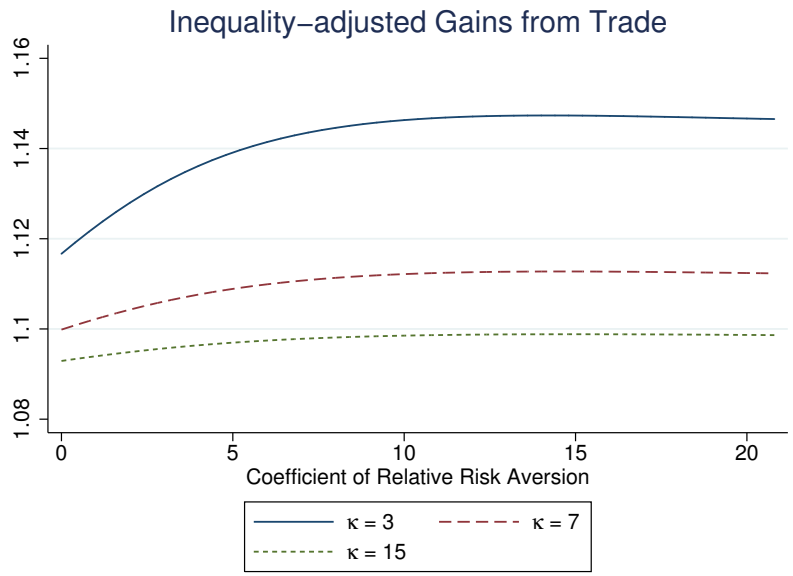
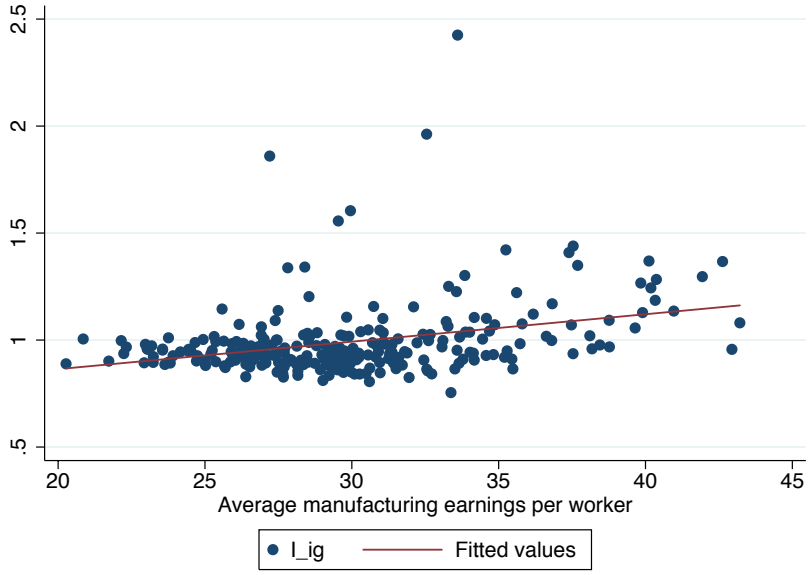


Figure 3.4: Relation between Import Competition and Earnings per Worker



we simulate counterfactual equilibria after an increase in China's technology level. Specifically, we study the effects of a sector-neutral productivity increase in China with $\hat{T}_{is}^{1/\theta} = 5$ for $i = China$ and all s.[36] We employ data from the World Input-Output Database (WIOD) for the year 2003 and focus on the manufacturing sector, as in the autarky exercise.[37]

As shown in Table **??**, the distributional effects of the productivity increase in China are strong for $\kappa \to 1$, where the coefficient of variation is almost equal to one. In this limit case there are also substantial outliers in terms of group-level gains. The dispersion of these gains falls quickly with κ. For instance, for $\kappa = 3$ the coefficient of variation has fallen to 0.35. Figure 3.5 visualizes these patterns.

[36]This counterfactual is closely related to the analysis in Hsieh and Ossa (2011), which examines how China's productivity growth affects worldwide real incomes. Hsieh and Ossa (2011) estimate annual sectoral productivity growth rates in China that range from 7.4% to 24.3%, with an average of 13.8%. The value of $\hat{T}_{is}^{1/\theta} = 5$ is on the high side of these estimates.

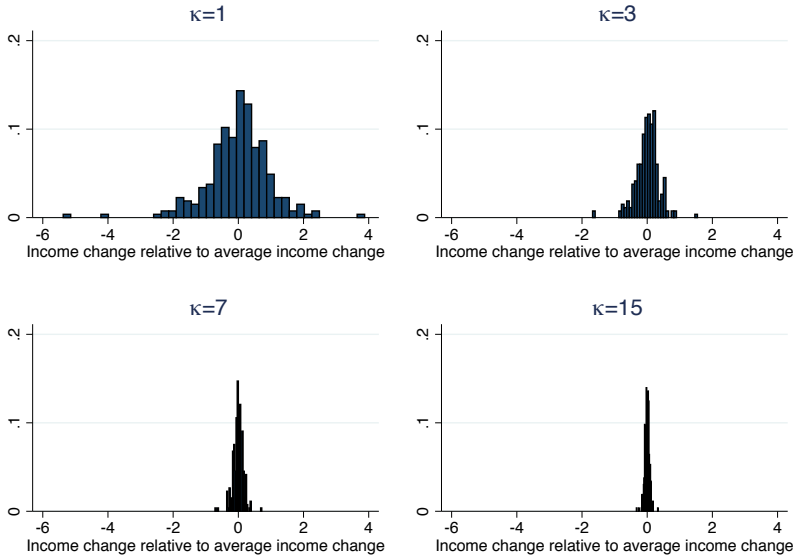
[37]The WIOD dataset is discussed in Timmer et al. (2015).

Table 3.7: $\hat{W}_{ig}$ in Germany - $\forall s : \hat{T}_{China,s}^{1/\theta} = 5$

	$\overline{\widehat{W}_{ig}}$ ($\Delta\%$)	$\widehat{W}_i$ ($\Delta\%$)	CV	Min.	Max.
$\kappa \to 1$ (Specific Factors)	0.609	0.647	0.966	-5.373	3.875
$\kappa = 3$	0.492	0.504	0.35	-1.675	1.542
$\kappa = 7$	0.459	0.465	0.154	-0.717	0.72
$\kappa = 15$	0.449	0.452	0.074	-0.332	0.337
$\kappa \to \infty$ (CDK)	0.44	0.44	0	1	1
N		265			

$\overline{\widehat{W}_{ig}} \equiv \frac{1}{G}\sum_g \widehat{W}_{ig} - 1$ and the first and second column display $100(\overline{\widehat{W}_{ig}} - 1), 100(\widehat{W}_i - 1)$ respectively. The third column displays the coefficient of variation (CV). For the final two columns: Min.= $[\min_g(\widehat{W}_{ig}) - 1]/[\overline{\widehat{W}_{ig}} - 1]$ and Max.=$[\max_g(\widehat{W}_{ig}) - 1]/[\overline{\widehat{W}_{ig}} - 1]$.

Figure 3.5: Distribution of Gains by Region - $\forall s : \hat{T}_{China,s}^{1/\theta} = 5$



A Bartik perspective: integrating simulations and data

The degree to which regions win or lose from the China shock depends on the change in their level of import-competition, as explained in section 3.2. Remember from Equation (3.23) that in the limit as $\kappa \to 1$, we have

$$\ln \hat{W}_{ig} = \ln \hat{W}_i + \ln \sum_s \pi_{igs} \hat{r}_{is}.$$

In Figure 3.6, we display this relationship between $\ln \hat{W}_{ig}$ and $\ln \sum_s \pi_{igs} \hat{r}_{is}$, the regional exposure to the national output response.[38] There is a linear and positive relationship between these two variables for each value of κ. For $\kappa \to 1$, this is in line with the theoretical result in equation (3.23) and Figure 3.6 suggests that this linearity persists for $\kappa > 1$, with a slope decreasing with κ. To examine this pattern more formally, we run the following regression for different values of κ:

$$\ln \hat{W}_{ig} = \ln \hat{W}_i + \beta \ln \sum_s \pi_{igs} \hat{r}_{is}.$$

This is the same regression as equation (3.24), which we employed earlier to analyze the relation between shocks to import competition and regional income changes in the German data. Now, we run this same regression, but on the simulated data, and obtain an estimate for β for different values of κ. Figure 3.7 plots the relation between the estimated β and the corresponding κ value, and demonstrates that β monotonically decreases with κ.

This monotone relationship between κ and β in the simulated data can now be used for a consistency check on our earlier empirical analysis. More specifically, we check if our estimated values for κ are consistent, from the model's point of view, with our estimates for β, which measure the impact of changes in import competition on real income. Remember that our estimates for κ in Table 3.4 ranged from 2.9 to 5. Figure 3.7 then indicates that these estimates are perfectly consistent with our estimates for β in Table 3.3, which are between 0.21 and 0.52.

This implies that, for a given increase in region-level import-competition, the model predictions on region-level income changes fit well with the observed changes in the data, given our estimate for κ. There is therefore a close connection between our theoretical framework and the patterns in the data.

[38] It is easy to show that if $\kappa \to 1$ then $\sum_s \pi_{igs} \hat{r}_{is} = 1/\hat{I}_{ig}$ (see Section 3.2). Here we use the expression $\sum_s \pi_{igs} \hat{r}_{is}$ because of its Bartik structure, which we will use to relate the model implications to those we see in the data.

Figure 3.6: Welfare-effects and changes in import-competition

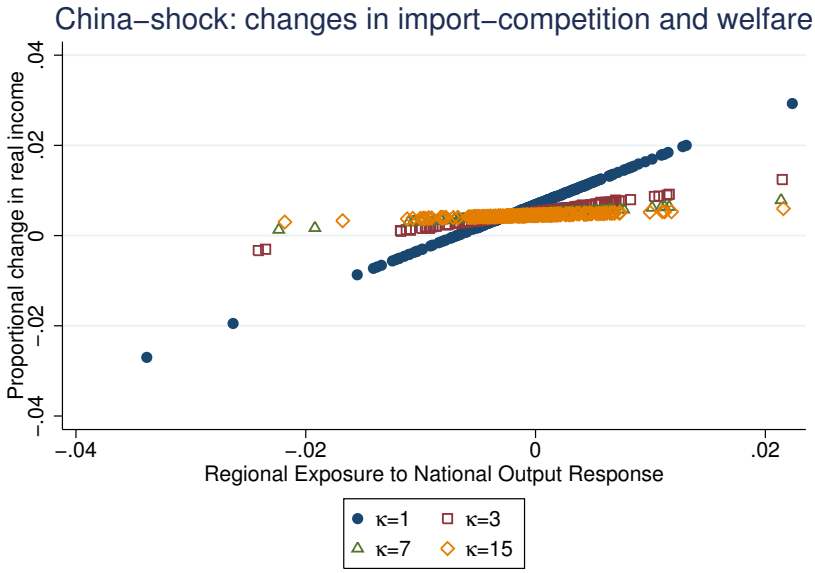
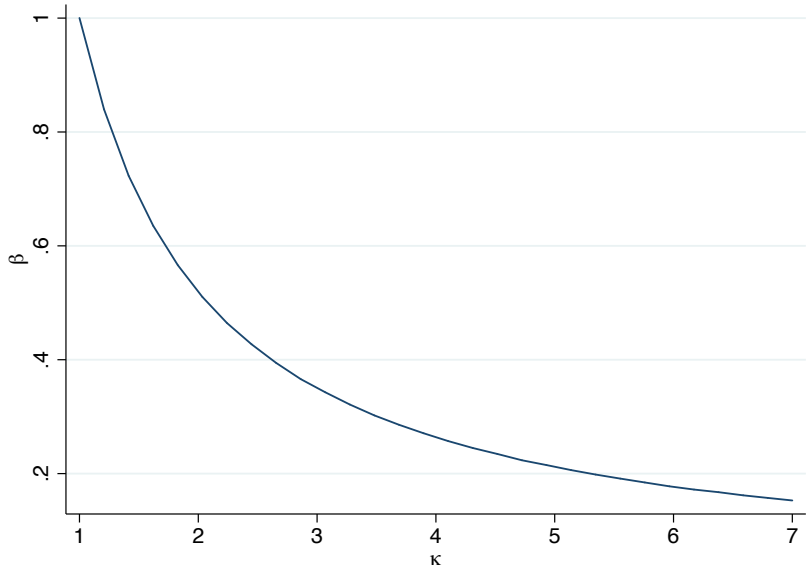


Figure 3.7: Relation between $\ln \sum_s \pi_{igs} \hat{r}_{is}$ and κ

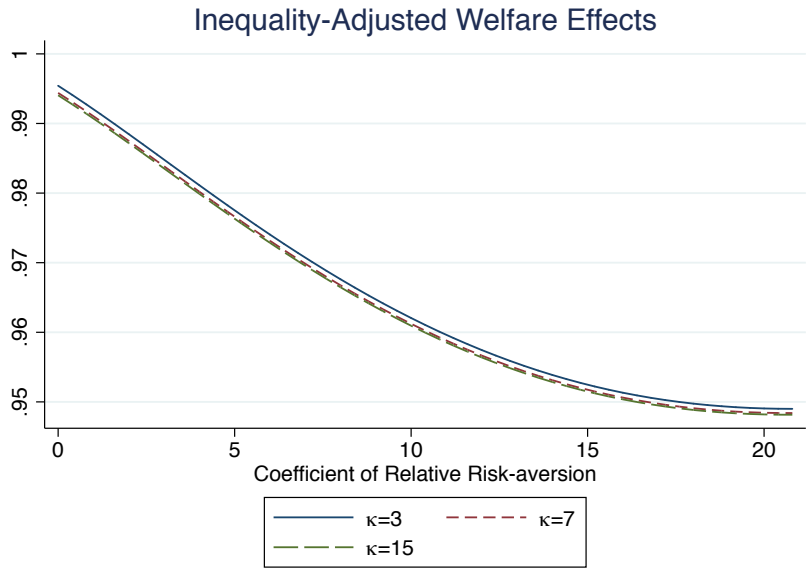


The coefficient β, on the vertical axis, is estimated in the following regressions: $\ln \hat{W}_{ig} = \ln \hat{W}_i + \beta \ln \sum_s \pi_{igs} \hat{r}_{is}$, which is run separately for different vectors of $\hat{r}_{is}$. Each vector of $\hat{r}_{is}$ is the outcome of a simulation under a different value of κ (horizontal axis).

Inequality-adjusted welfare

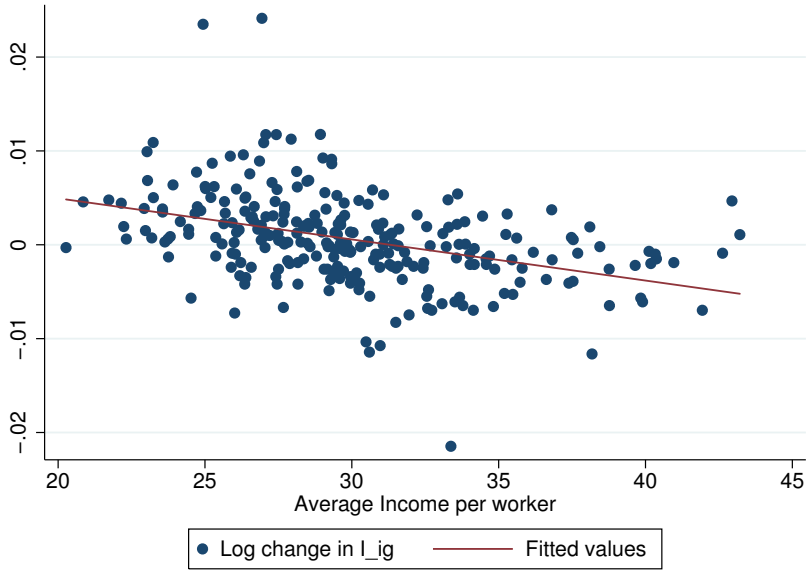
We now use the distribution of group-level welfare effects from the China shock to compute the inequality-adjusted welfare effect of this shock for Germany. In Figure 3.8 we plot the inequality-adjusted welfare effect, $\hat{U}_i$ for $i = Germany$. By definition, this is equal to the standard aggregate effect ($\hat{W}_i$) when the coefficient of inequality aversion (ρ) is zero. The figure reveals that the inequality-adjusted welfare gain is decreasing in ρ, so that for any positive level of ρ we have $\hat{U}_i > \hat{W}_i$. The reason for this is that, as shown in Figure 3.9, there is a negative covariance between the change in the degree of import competition ($\hat{I}_{ig}$) and the initial income level (Y_{ig}).[39] This implies that the cross-region distributional impact of the China shock is pro-rich.

Figure 3.8: Inequality-Adjusted welfare-effects from the China shock



[39] As mentioned in the previous footnote, in the limit $\kappa \to 1$ we have $\sum_s \pi_{igs}\hat{r}_{is} = 1/\hat{I}_{ig}$, hence $\ln \hat{W}_{ig} = \ln \hat{W}_i - \ln \hat{I}_{ig}$.

Figure 3.9: Relation between income per worker and $\ln \hat{I}_g$ - China shock



3.6 Conclusion

This paper develops and applies a framework to quantify the effect of trade on aggregate welfare as well as the distribution of this aggregate effect across different groups of workers. The framework combines a multi-sector gravity model of trade with a Roy-type model of the allocation of workers across sectors. By opening to trade, a country gains in the aggregate by specializing according to its comparative advantage, but the distribution of these gains is unequal as labor demand increases (decreases) for groups of workers specialized in export-oriented (import-oriented) sectors. The model generalizes the specific-factors intuition to a setting with labor reallocation, while maintaining analytical tractability for any number of groups and countries. Our new notion of "inequality-adjusted" welfare effect of trade captures the full cross-group distribution of welfare changes in one measure, as the counterfactual scenario is evaluated by a risk-averse agent behind the veil of ignorance regarding the group to which she belongs. The quantitative application uses trade and labor allocation data across regions in the US and Germany to compute the aggregate and distributional effects of a shock to trade costs or foreign technology levels. For the extreme case in which the country moves back to autarky we find that inequality-adjusted gains from trade are larger than the aggregate gains for both countries, as between-group inequality falls with trade relative to autarky, but the opposite happens for the shock in which China expands in the world economy.

Bibliography

Aghion, Philippe, Ufuk Akcigit, and Peter Howitt (2013). *What do we learn from Schumpeterian growth theory?* Tech. rep. National Bureau of Economic Research.

Aghion, Philippe, Nick Bloom, Richard Blundell, Rachel Griffith, and Peter Howitt (2005). "Competition and Innovation: An Inverted-U Relationship". In: *The Quarterly Journal of Economics*, pp. 701–728.

Aguiar, Mark, Erik Hurst, and Loukas Karabarbounis (2013). "Time use during the great recession". In: *The American Economic Review* 103.5, pp. 1664–1696.

Akcigit, Ufuk, Harun Alp, and Michael Peters (2014). "Lack of Selection and Imperfect Managerial Contracts: Firm Dynamics in Developing Countries". In: *working paper, University of Pennsylvania*.

Allcott, Hunt, Allan Collard-Wexler, and Stephen D O'Connell (2014). "How do electricity shortages affect productivity? Evidence from India". In: *NBER working paper*.

Antras, Pol, Alonso de Gortari, and Oleg Itskhoki (2016). *Inequality, costly redistribution and welfare in an open economy*. Tech. rep. Working Paper.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare (2012). "New Trade Models, Same Old Gains?" In: *American Economic Review* 102.1, pp. 94–130.

Artuç, Erhan, Shubham Chaudhuri, and John McLaren (2010). "Trade Shocks and Labor Adjustment: A Structural Empirical Approach". In: *American Economic Review* 100.3, pp. 1008–45.

Asker, John, Allan Collard-Wexler, and Jan De Loecker (2014). "Dynamic inputs and resource (mis) allocation". In: *Journal of Political Economy* 122.5, pp. 1013–1063.

Atkeson, Andrew and Ariel Burstein (2008). "Pricing-to-market, trade costs, and international relative prices". In: *The American Economic Review* 98.5, pp. 1998–2031.

Atkin, David and Dave Donaldson (2014). "Who is Getting Globalized? The Size and Nature of Intranational Trade Costs". In: *UCLA and Stanford unpublished mimeo*.

Autor, David H., David Dorn, and Gordon H. Hanson (2013). "The China Syndrome: Local Labor Market Effects of Import Competition in the United States". In: *American Economic Review* 103.6, pp. 2121–68.

Autor, David, David Dorn, Gordon H Hanson, and Jae Song (2014). "Trade Adjustment: Worker-Level Evidence". In: *The Quarterly Journal of Economics* 129.4, pp. 1799–1860.

Banerjee, Abhijit V and Esther Duflo (2014). "Do firms want to borrow more? Testing credit constraints using a directed lending program". In: *The Review of Economic Studies* 81.2, pp. 572–607.

Bartik, Timothy J (1991). "Boon or boondoggle? The debate over state and local economic development policies". In:

Bénassy, Jean-Pascal (1996). "Taste for variety and optimum production patterns in monopolistic competition". In: *Economics Letters* 52.1, pp. 41–47.

Blanchard, Olivier Jean and Nobuhiro Kiyotaki (1987). "Monopolistic competition and the effects of aggregate demand". In: *The American Economic Review*, pp. 647–666.

Bollard, Albert, Peter J Klenow, and Gunjan Sharma (2013). "India's Mysterious Manufacturing Miracle". In: *Review of Economic Dynamics* 16.1, pp. 59–85.

Buera, Francisco J, Joseph P Kaboski, and Yongseok Shin (2015). *Entrepreneurship and Financial Frictions: A Macro-Development Perspective*. Tech. rep. National Bureau of Economic Research.

Burstein, Ariel, Eduardo Morales, and Jonathan Vogel (2015). "Accounting for changes in between-group inequality". In: *unpublished manuscript, UCLA, Princeton and NYU*.

Burstein, Ariel and Jonathan Vogel (2012). "International trade, technology, and the skill premium". In: *Manuscript, Columbia University and UCLA*.

Caliendo, Lorenzo and Fernando Parro (2014). "Estimates of the Trade and Welfare Effects of NAFTA". In: *The Review of Economic Studies, forthcoming*.

Chaney, Thomas (2008). "Distorted gravity: The intensive and extensive margins of international trade". In: *The American Economic Review* 98.4, pp. 1707–1721.

Cockburn, John, Bernard Decaluwé, and Véronique Robichaud (2008). "Trade Liberalization and Poverty: A CGE Analysis of the 1990s Experience in Africa and Asia". In:

Cordoba, Juan Carlos and Genevieve Verdier (2008). "Inequality and growth: Some welfare calculations". In: *Journal of Economic Dynamics and Control* 32.6, pp. 1812–1829.

Coşar, A Kerem (2013). "Adjusting to Trade Liberalization: Reallocation and Labor Market Policies". In:

Coşar, A. Kerem, Nezih Guner, and James Tybout (2013). *Firm Dynamics, Job Turnover, and Wage Distributions in an Open Economy*. Working Papers 732. Barcelona Graduate School of Economics.

Costinot, Arnaud, Dave Donaldson, and Ivana Komunjer (2012). "What Goods Do Countries Trade? A Quantitative Exploration of Ricardo's Ideas". In: *The Review of economic studies* 79.2, pp. 581–608.

Costinot, Arnaud and Andrés Rodríguez-Clare (2014). "Trade Theory with Numbers: Quantifying the Consequences of Globalization". In: *Handbook of International Economics* 4, p. 197.

Costinot, Arnaud and Jonathan Vogel (2010). "Matching and Inequality in the World Economy". In: *Journal of Political Economy* 118.4.

— (2014). *Beyond Ricardo: Assignment Models in International Trade*. Tech. rep. National Bureau of Economic Research.

Dauth, Wolfgang, Sebastian Findeisen, and Jens Suedekum (2014). "The rise of the East and the Far East: German labor markets and trade integration". In: *Journal of the European Economic Association* 12.6, pp. 1643–1675.

De Loecker, Jan and Frederic Warzynski (2012). "Markups and Firm-Level Export Status". In: *The American Economic Review* 102.6, pp. 2437–2471.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum (2008). "Global Rebalancing with Gravity: Measuring the Burden of Adjustment". In: *IMF Staff Papers* 55.3, pp. 511–540.

Dix-Carneiro, Rafael (2014). "Trade liberalization and labor market dynamics". In: *Econometrica* 82.3, pp. 825–885.

Dix-Carneiro, Rafael and Brian Kovak (2014). *Trade Reform and Regional Dynamics: Evidence from 25 Years of Brazilian Matched Employer-Employee Data*. Tech. rep. Working Paper.

Eaton, Jonathan and Samuel Kortum (2002). "Technology, Geography, and Trade". In: *Econometrica* 70.5, pp. 1741–1779.

Epifani, Paolo and Gino Gancia (2011). "Trade, markup heterogeneity and misallocations". In: *Journal of International Economics* 83.1, pp. 1–13.

Evans, David S (1987). "Tests of alternative theories of firm growth". In: *The Journal of Political Economy*, pp. 657–674.

Faber, Benjamin (2014). "Trade Liberalization, the Price of Quality, and Inequality: Evidence from Mexican Store Prices". In:

Fajgelbaum, Pablo D and Amit K Khandelwal (2014). "Measuring the unequal gains from trade". In: *unpublished manuscript, Colombia and UCLA*.

García-Santana, Manuel and Josep Pijoan-Mas (2014). "The reservation laws in India and the misallocation of production factors". In: *Journal of Monetary Economics* 66, pp. 193–209.

Geurts, Karen and Johannes Van Biesebroeck (2014). "Job creation, firm creation, and de novo entry". In: *CEPR Discussion Paper No. DP10118*.

Gilbert, Richard (2006). "Looking for Mr. Schumpeter: Where are we in the competition-innovation debate?" In: *Innovation Policy and the Economy, Volume 6*. The MIT Press, pp. 159–215.

Gourinchas, Pierre-Olivier (1999). "Exchange rates do matter: French job reallocation and exchange rate turbulence, 1984-1992". In: *European Economic Review* 43.7, pp. 1279–1316.

Hall, Robert (1986). "Market structure and macroeconomic fluctuations". In: *Brookings papers on economic activity* 1986.2, pp. 285–338.

Haltiwanger, John, Ron S Jarmin, and Javier Miranda (2013). "Who creates jobs? Small versus large versus young". In: *Review of Economics and Statistics* 95.2, pp. 347–361.

Head, Keith and Thierry Mayer (2014). "Gravity Equations: Workhorse, Toolkit, and Cookbook". In: *Handbook of International Economics*. Ed. by Kenneth Rogoff Elhanan Helpman and Gita Gopinath. Vol. 4. Handbook of International Economics. Elsevier, pp. 131–195.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante (2008). "Insurance and opportunities: A welfare analysis of labor market risk". In: *Journal of Monetary Economics* 55.3, pp. 501–525.

Helpman, Elhanan, Oleg Itskhoki, Marc-Andreas Muendler, and Stephen J. Redding (2012). *Trade and Inequality: From Theory to Estimation*. NBER Working Papers 17991. National Bureau of Economic Research, Inc.

Hsieh, Chang-Tai, Erik Hurst, Charles I. Jones, and Peter J. Klenow (2013). *The Allocation of Talent and U.S. Economic Growth*. NBER Working Papers 18693. National Bureau of Economic Research, Inc.

Hsieh, Chang-Tai and Peter J Klenow (2009). "Misallocation and Manufacturing TFP in China and India." In: *Quarterly Journal of Economics* 124.4.

— (2014). "The Life Cycle of Plants in India and Mexico". In: *The Quarterly Journal of Economics* 129.3, pp. 1035–1084.

Hsieh, Chang-Tai and Ralph Ossa (2011). *A global view of productivity growth in China*. Tech. rep. National Bureau of Economic Research.

Itskhoki, Oleg and Benjamin Moll (2015). "Optimal development policies with financial frictions". In: *working paper, Princeton University*.

Jaimovich, Nir (2007). "Firm dynamics and markup variations: Implications for sunspot equilibria and endogenous economic fluctuations". In: *Journal of Economic Theory* 137.1, pp. 300–325.

Jones, Charles and Peter Klenow (2016). "Beyond GDP? Welfare across Countries and Time". In: *AER, forthcoming*.

Kambourov, Gueorgui (2009). "Labour market regulations and the sectoral reallocation of workers: The case of trade reforms". In: *The Review of Economic Studies* 76.4, pp. 1321–1358.

Kline, Patrick (2008). "Understanding Sectoral Labor Market Dynamics: An Equilibrium Analysis of the Oil and Gas Field Services Industry". In: *Cowles Foundation Discussion Paper*.

Kovak, Brian K (2013). "Regional Effects of Trade Reform: What is the Correct Measure of Liberalization?" In: *The American Economic Review* 103.5, pp. 1960–1976.

Krishna, Pravin, Jennifer P. Poole, and Mine Zeynep Senses (2012). "Trade, Labor Market Frictions, and Residual Wage Inequality across Worker Groups". In: *American Economic Review* 102.3, pp. 417–23.

Krugman, Paul (1980). "Scale economies, product differentiation, and the pattern of trade". In: *The American Economic Review* 70.5, pp. 950–959.

Kuntchev, Veselin, Rita Ramalho, Jorge Rodríguez-Meza, and Judy S Yang (2014). "What have we learned from the enterprise surveys regarding access to credit by SMEs?" In: *World Bank Policy Research Working Paper* 6670.

Lagakos, David and Michael E Waugh (2013). "Selection, Agriculture, and Cross-Country Productivity Differences". In: *The American Economic Review* 103.2, pp. 948–980.

Macchiavello, Rocco and Ameet Morjaria (2015). *Competition and relational contracts: Evidence from Rwanda's mills*.

Martin, Leslie A, Shanthi Nataraj, and Ann Harrison (2014). "In with the Big, Out with the Small: Removing Small-Scale Reservations in India". In: *NBER working paper*.

McLaren, John and Shushanik Hakobyan (2010). *Looking for local labor market effects of NAFTA*. Tech. rep. National Bureau of Economic Research.

Melitz, Marc J (2003). "The impact of trade on intra-industry reallocations and aggregate industry productivity". In: *Econometrica* 71.6, pp. 1695–1725.

Menezes-Filho, Naercio Aquino and Marc-Andreas Muendler (2011). *Labor Reallocation in Response to Trade Reform*. NBER Working Papers 17372. National Bureau of Economic Research, Inc.

Midrigan, Virgiliu and Daniel Yi Xu (2014). "Finance and Misallocation: Evidence from Plant-Level Data". In: *American Economic Review* 104.2, pp. 422–458.

Moll, Benjamin (2014). "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?" In: *The American Economic Review* 104.10, pp. 3186–3221.

Peters, Michael (2013). "Heterogeneous Mark-ups, Growth and Endogenous Misallocation". In: *mimeo, Yale University*.

Porto, Guido G (2006). "Using survey data to assess the distributional effects of trade policy". In: *Journal of International Economics* 70.1, pp. 140–160.

Rajan, Raghuram and Luigi Zingales (1998). "Financial Dependence and Growth". In: *The American Economic Review* 88.3, pp. 559–586.

Restuccia, Diego and Richard Rogerson (2008). "Policy distortions and aggregate productivity with heterogeneous establishments". In: *Review of Economic Dynamics* 11.4, pp. 707–720.

— (2013). "Misallocation and productivity". In: *Review of Economic Dynamics* 16.1, pp. 1–10.

Revenga, Ana L (1992). "Exporting jobs?: The impact of import competition on employment and wages in US manufacturing". In: *The Quarterly Journal of Economics*, pp. 255–284.

Ritter, Moritz (2012). *Trade and Inequality in a Directed Search Model with Firm and Worker Heterogeneity*. Tech. rep. Department of Economics, Temple University.

Roy, Andrew Donald (1951). "Some thoughts on the distribution of earnings". In: *Oxford economic papers* 3.2, pp. 135–146.

Savard, Luc (2003). "Poverty and Income Distribution in a CGE-Household Micro-Simulation Model: Top-Down/Bottom Up Approach". In:

Schaumans, Catherine and Frank Verboven (2015). "Entry and competition in differentiated products markets". In: *Review of Economics and Statistics* 97.1, pp. 195–209.

Tewari, Ishani and Joshua Wilde (2014). "Multiproduct Firms, Product Scope and Productivity: Evidence from India's Product Reservation Policy". In: *working paper, University of South Florida*.

Timmer, Marcel P, Erik Dietzenbacher, Bart Los, Robert Stehrer, and Gaaitzen J Vries (2015). "An Illustrated User Guide to the World Input–Output Database: the Case of Global Automotive Production". In: *Review of International Economics*.

Topalova, Petia (2010). "Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India". In: *American Economic Journal: Applied Economics* 2.4, pp. 1–41.

Young, Alwyn (2014). "Structural Transformation, the Mismeasurement of Productivity Growth, and the Cost Disease of Services". In: *American Economic Review* 104.11, pp. 3635–67.

Appendix A

Competition, Financial Constraints and Misallocation

A.1 Labor market equilibrium

Expressions for output and TFP

We can express each firm's capital as a share of aggregate capital. To that end, we rewrite capital demand for constrained and unconstrained firms as:

$$k_{it} = \mu_{it}^{\frac{1}{\eta-1}} a_{it}^{\frac{\eta}{1-\eta}} \frac{Q_t}{M} \left(\frac{P_t(1-\alpha)}{w_t}\right)^{\frac{\eta-\alpha\eta}{1-\eta}} \left(\frac{\alpha}{\omega_{it}}\right)^{\frac{1+\alpha\eta-\eta}{1-\eta}}$$

where $\omega_{it} = r_{it}$ if the firm is unconstrained, and $\omega_{it} > r_{it}$ otherwise. Writing k_{it} as a fraction of aggregate capital, we find:

$$k_{it} = \frac{\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}}}{\sum_{i=1}^{M}\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}}} K_t$$

Similarly, for labor, starting from the labor demand equation $l_{it} = \left(\frac{(1-\alpha)}{\mu_{it}}\frac{P_t}{w_t}\left(\frac{Q_t}{M}\right)^{1-\eta} a_{it}^{\eta} k_{it}^{\alpha\eta}\right)^{\frac{1}{1+\alpha\eta-\eta}}$

$$l_{it} = \frac{\left(\frac{a_{it}^{\eta} k_{it}^{\alpha\eta}}{\mu_{it}}\right)^{\frac{1}{1+\alpha\eta-\eta}}}{\sum_{i=1}^{M}\left(\frac{a_{it}^{\eta} k_{it}^{\alpha\eta}}{\mu_{it}}\right)^{\frac{1}{1+\alpha\eta-\eta}}} L$$

Plugging in the value for k_{it}

$$l_{it} = \frac{\frac{\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{\alpha\eta}}\right)^{\frac{1}{1-\eta}}}{\left[\sum_{i=1}^{M}\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}}\right]^{\frac{\alpha\eta}{1+\alpha\eta-\eta}}}}{\sum_{i=1}^{M}\frac{\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{\alpha\eta}}\right)^{\frac{1}{1-\eta}}}{\left[\sum_{i=1}^{M}\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}}\right]^{\frac{\alpha\eta}{1+\alpha\eta-\eta}}}} L$$

The expressions for k_{it}, l_{it} can then be used to find an expression for the composite good: $Q_t = M^{1-\frac{1}{\eta}}\left[\sum_{i=1}^{M}(y_{it})^{\eta}\right]^{\frac{1}{\eta}} = M^{1-\frac{1}{\eta}}\left[\sum_{i=1}^{M}\left(a_{it}k_{it}^{\alpha}l_{it}^{1-\alpha}\right)^{\eta}\right]^{\frac{1}{\eta}}$.

$$Q_t = MK_t^{\alpha}L^{1-\alpha}\left[E_{it}\left(a_{it}^{\eta}\left(\frac{\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}}}{\sum_{i=1}^{M}\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}}}\right)^{\alpha\eta}\left(\frac{\frac{\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{\alpha\eta}}\right)^{\frac{1}{1-\eta}}}{\left[\sum_{i=1}^{M}\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}}\right]^{\frac{\alpha\eta}{1+\alpha\eta-\eta}}}}{\sum_{i=1}^{M}\frac{\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{\alpha\eta}}\right)^{\frac{1}{1-\eta}}}{\left[\sum_{i=1}^{M}\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}}\right]^{\frac{\alpha\eta}{1+\alpha\eta-\eta}}}}\right)^{\eta-\alpha\eta}\right)\right]^{1/\eta}$$

Therefore:

$$Q_t = TFP_t K_t^{\alpha}L^{1-\alpha} \tag{A.1}$$

where

$$TFP_t \equiv M\left[E_{it}\left(a_{it}^{\eta}\left(\frac{\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}}}{\sum_{i=1}^{M}\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}}}\right)^{\alpha\eta}\left(\frac{\frac{\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{\alpha\eta}}\right)^{\frac{1}{1-\eta}}}{\left[\sum_{i=1}^{M}\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}}\right]^{\frac{\alpha\eta}{1+\alpha\eta-\eta}}}}{\sum_{i=1}^{M}\frac{\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{\alpha\eta}}\right)^{\frac{1}{1-\eta}}}{\left[\sum_{i=1}^{M}\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}}\right)^{\frac{1}{1-\eta}}\right]^{\frac{\alpha\eta}{1+\alpha\eta-\eta}}}}\right)^{\eta-\alpha\eta}\right)\right]^{1/\eta} \tag{A.2}$$

Labor market equilibrium

$$L = \sum_{i=1}^{M} \left(\frac{(1-\alpha)}{\mu_{it}} \frac{P_t}{w_t} \left(\frac{Q_t}{M} \right)^{1-\eta} a_{it}^{\eta} k_{it}^{\alpha\eta} \right)^{\frac{1}{1+\alpha\eta-\eta}}$$

$$L = \left((1-\alpha) \frac{P_t}{w_t} \left(\frac{Q_t}{M} \right)^{1-\eta} \right)^{\frac{1}{1+\alpha\eta-\eta}} \sum_{i=1}^{M} \left(\frac{a_{it}^{\eta}}{\mu_{it}} k_{it}^{\alpha\eta} \right)^{\frac{1}{1+\alpha\eta-\eta}}$$

$$L = \left((1-\alpha) \frac{P_t}{w_t} \left(\frac{TFP_t K_t^{\alpha} L^{1-\alpha}}{M} \right)^{1-\eta} \right)^{\frac{1}{1+\alpha\eta-\eta}} K_t^{\frac{\alpha\eta}{1+\alpha\eta-\eta}} \sum_{i=1}^{M} \left(\frac{a_{it}^{\eta}}{\mu_{it}} \left(\frac{\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}} \right)^{\frac{1}{1-\eta}}}{\sum_{i=1}^{M} \left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}} \right)^{\frac{1}{1-\eta}}} \right)^{\alpha\eta} \right)^{\frac{1}{1+\alpha\eta-\eta}}$$

$$L^{\frac{\alpha}{1+\alpha\eta-\eta}} = \left((1-\alpha) \frac{P_t}{w_t} \left(\frac{TFP_t}{M} \right)^{1-\eta} \right)^{\frac{1}{1+\alpha\eta-\eta}} K_t^{\frac{\alpha}{1+\alpha\eta-\eta}} \sum_{i=1}^{M} \left(\frac{a_{it}^{\eta}}{\mu_{it}} \left(\frac{\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}} \right)^{\frac{1}{1-\eta}}}{\sum_{i=1}^{M} \left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}} \right)^{\frac{1}{1-\eta}}} \right)^{\alpha\eta} \right)^{\frac{1}{1+\alpha\eta-\eta}}$$

$$\frac{P_t}{w_t} = \left(\frac{L}{K_t} \right)^{\alpha} \frac{1}{(1-\alpha)\left(\frac{TFP_t}{M}\right)^{1-\eta}} \left[\sum_{i=1}^{M} \left(\frac{a_{it}^{\eta}}{\mu_{it}} \left(\frac{\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}} \right)^{\frac{1}{1-\eta}}}{\sum_{i=1}^{M} \left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}} \right)^{\frac{1}{1-\eta}}} \right)^{\alpha\eta} \right)^{\frac{1}{1+\alpha\eta-\eta}} \right]^{\eta-\alpha\eta-1}$$

So $\frac{P_t}{w_t}$ is decreasing in TFP and in $\left[\sum_{i=1}^{M} \left(\frac{a_{it}^{\eta}}{\mu_{it}} \left(\frac{\left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}} \right)^{\frac{1}{1-\eta}}}{\sum_{i=1}^{M} \left(\frac{a_{it}^{\eta}}{\mu_{it}\omega_{it}^{1+\alpha\eta-\eta}} \right)^{\frac{1}{1-\eta}}} \right)^{\alpha\eta} \right)^{\frac{1}{1+\alpha\eta-\eta}} \right]$.

A.2 Model with young firms

Agents The worker side of the model is unaltered from the baseline model. On the firm side, there continues to be an exogenous, finite set M of firm-owners. In this version of the model, heterogeneity across firms arises from the date at which they are born. Before the start of each period, qM new firms are born with capital levels $k_0 \equiv \zeta \frac{K}{M}$, where K is aggregate capital and $0 < \zeta < 1$. At the same time, a set of firms qM dies before the start of the period, such that the total number of firms remains constant.[1]

[1]The ex-ante probability that any firm dies is constant at q, but this probability is not independent across firms as I assume that each period the dying firms hold the same fraction of aggregate capital.

Firm-owner i has the following intertemporal preferences at time t:

$$U_{it} = \sum_{s=t}^{\infty} (q\beta)^{s-t} d_{is}$$

Where β is the discount factor, q is the ex-ante probability a firm dies in any given period and d_{it} is firm-owner consumption.

Production of varieties Each firm produces a variety i with a Cobb-Douglas production function, using capital k_{it} and labor l_{it} as inputs. There is no variation in productivity across firms.

$$y_{it} = k_{it}^{\alpha} l_{it}^{1-\alpha} \tag{A.3}$$

Investment $k_{it+1} = x_{it} + (1-\delta)k_{it}$ is modeled exactly as in the baseline model. The same holds for the definition of the final good, firm-level demand (1.3), the price index (1.4), the budget constraint (1.5), and the financial constraint (1.6).

Market structure and optimization in steady state

The market structure and firm-problem are equivalent to the set-up in the baseline model, except that there is no firm-level productivity volatility to be taken into account. Since firms play a one-period game of quantity competition, each firm i sets a quantity y_{it+1} for sale, conditional on the quantities chosen by the other firms in the economy. As discussed in the previous subsection, firms make decisions about l_{it+1}, k_{it+1} in period t, given the budget constraint $P_t(k_{it+1}+d_{it}) \leq z_{it}$. Therefore, any firm i's optimal decisions are $k_{it+1}\left(z_{it}, \mathbf{y_{-it+1}}\right)$, $l_{it+1}\left(z_{it}, \mathbf{y_{-it+1}}\right)$, where (z_{it}) characterizes the state for firm i and $\mathbf{y_{-it+1}}$ is the vector of decisions on y_{jt+1} for all $j \neq i$. Through the production function (A.3), the choice of k_{it+1}, l_{it+1} determines y_{it+1} and thereby $p_{it+1}(y_{it+1}, \mathbf{y_{-it+1}})$ as firms incorporate the demand function (1.3) into their optimization. As such, this setting entails the following intertemporal problem for the firm, where $\pi_{it}(k_{it}, l_{it}, \mathbf{y_{-it}}) \equiv p_{it}(y_{it}, \mathbf{y_{-it}})y_{it} - w_t l_{it}$:

$$\begin{aligned} \max_{d_{it}, k_{it+1}, l_{it+1}} \mathcal{L} = & \sum_{t=s}^{\infty} E_s\left[\beta^{t-s} d_{it}\right] + \\ & \sum_{t=s}^{\infty} E_s\left[\lambda_{it}\left(\pi_{it}(k_{it}, l_{it}, \mathbf{y_{-it}}) + P_t\left[(1-\delta)k_{it} - k_{it+1} - d_{it}\right]\right) + \Phi_{it}(d_{it})\right] \end{aligned} \tag{A.4}$$

Since each firm's decision on y_{it+1} depends on $(z_{it}, \mathbf{y_{-it+1}})$, $\mathbf{y_{it+1}}$ will be determined by $F(z(t))$, the distribution of z_{it}, and by the conditions in the labor and goods market implied by M, L.

$$\begin{aligned} & k_{it+1}\left(z_{it}, F(z(t)), M, L\right) \\ & l_{it+1}\left(z_{it}, F(z(t)), M, L\right) \end{aligned} \tag{A.5}$$

From here on, the optimization is exactly as in the baseline model, with equivalent expressions for the demand elasticity, the labor choice and the capital choice.

Steady state equilibrium

An *equilibrium* consists of a set of prices P_t, w_t, p_{it}, a set of consumption $d_{it}(z_{it}, F(z(t)))$, capital $k_{it+1}(z_{it}, F(z(t)))$ and labor $l_{it}(z_{it-1}, F(z(t-1)))$ decisions by firm-owners and consumption by workers $\frac{w_t}{P_t}L$ that satisfy

- the labor market clearing condition

$$L = \sum_{i=1}^{M} l_{it} \tag{A.6}$$

- the goods market clearing condition

$$Q_t = \sum_{i=1}^{M}(x_{it} + d_{it}) + \int_{l \in L} c_{lt} dl \tag{A.7}$$

- the optimality conditions for labor and capital for each firm i, conditional on the choices of l_{jt}, k_{jt} of all firms $j \neq i$.
- market-clearing for each variety i: $y_{it} = q_{it}$, satisfying the expression for firm demand.
- the equalized budget constraint $P_t(k_{it+1}+d_{it}) = z_{it}$, and the financial constraint $d_{it} \geq 0$.
- Firms are born with a capital level k_0. This capital level k_0, with $k_0 = \zeta k^*$, where k_0 is inherited from the dead firms, such that necessarily :$\bar{k} \geq k_0$. And here, $\bar{k} = \frac{K}{M}$. In steady state, we know that $k_0 < k_1$ (i.e. since K is constant, all firms are born with the same k_0 and afterwards grow their capital.

Steady state conditions

- $K_t = K$
- $P_t/w_t = P/w$
- $F(z(t)) = F(z)$,

An implication of $K_t = K$ is that capital growth by surviving firms will have to equal the capital loss from firms dying.

Labor and capital decisions in steady state

It will again be convenient to characterize the solution to the firm's optimization problem by taking the perspective of the cost-minimization problem given the optimal markup characterized in (1.9). The cost-minimization problem implies the following optimal labor demand in steady state:

$$l_{it} = \left(\frac{(1-\alpha)}{\mu_{it}} \frac{P}{w} \left(\frac{Q}{M} \right)^{1-\eta} k_{it}^{\alpha\eta} \right)^{\frac{1}{1+\alpha\eta-\eta}} \tag{A.8}$$

There are two cases for the firm's capital choice: either $\Phi_{it} = 0$, or $\Phi_{it} > 0$.

Unconstrained firms First consider the case where a firm has $\Phi_{it} = 0$.

$$k^* = \mu_U^{\frac{1}{\eta-1}} \frac{Q}{M} \left(\frac{P(1-\alpha)}{w} \right)^{\frac{\eta-\alpha\eta}{1-\eta}} \left(\frac{\alpha}{r_{it}} \right)^{\frac{1+\alpha\eta-\eta}{1-\eta}} \tag{A.9}$$

with the new definition $r_{it} \equiv \left(\frac{1}{q\beta} + \delta - 1 \right)$ and μ_U the markup of the unconstrained firm.

Constrained firms When the financial constraint binds, i.e. $\Phi_{it} > 0$. Capital grows as allowed by the budget constraint

$$k_{it+1} = (1-\delta)k_{it} + \left(\left(\frac{(1-\alpha)}{\mu_{it}} \right)^{\frac{\eta-\alpha\eta}{1+\alpha\eta-\eta}} - \left(\frac{(1-\alpha)}{\mu_{it}} \right)^{\frac{1}{1+\alpha\eta-\eta}} \right) \left(\frac{P}{w} \left(\frac{Q}{M} \right)^{1-\eta} k_{it}^{\alpha\eta} \right)^{\frac{1}{1+\alpha\eta-\eta}} \tag{A.10}$$

This will then imply the following lemma for the capital distribution $H(k)$ in steady state, where τ is the number of periods since the firm was born:

Lemma 5. *Steady state H(k) is given by:*

- *When* $\Phi_\tau = 0$*, then* $k_{i\tau} = k^*$
- *When* $\Phi_\tau > 0$*, then* $k_{i\tau} = G_\tau k_0$*, where* $G_\tau = \Pi_{s=0}^{\tau-1}(1+g_s)$ *and* $g_s = \frac{k_{is+1}}{k_{is}}$

This way, the capital distribution in this economy is essentially isomorphic to the distribution of the baseline model. Furthermore, the other elements of the system of equations - the markup distribution, TFP, K, $\frac{P}{w}, \Omega$ - are isomorphic as well, after properly adjusting for the constant productivity. Therefore, this model exhibits analogous comparative statics on M as the baseline model.

A.3 Markup Measurement

The markup measurement is based on De Loecker and Warzynski (2012), who elaborated on the framework introduced by Hall (1986). The main structural assumption for this markup measurement is cost-minimization by firms. Therefore, setup the Lagrangian for cost-minimization on the variable inputs $X_{it}^1, ..., X_{it}^V$,:

$$L_{it}(X_{it}^1, ..., X_{it}^V, K_{it}) = \sum_{v=1}^{V} P^{X_{it}^v} X_{it}^v + r_{it} K_{it} + \lambda_{it}(Q_{it} - Q_{it}(X_{it}^1, ..., X_{it}^V, K_{it}))$$

$$\text{FOC}: \frac{\partial L_{it}}{\partial X_{it}^v} = P^{X_{it}^v} - \lambda_{it} \frac{\partial Q_{it}(.)}{\partial X_{it}} = 0 \Rightarrow \frac{P_{it}^Y}{\lambda_{it}} = \frac{\partial Q_{it}(.) X_{it}^v}{\partial X_{it} Y_{it}} \frac{P_{it}^Y Y_{it}}{P^{X_{it}^v} X_{it}^v}$$

Which implies:

$$\mu_{it} = \frac{\theta_{it}^{X^v}}{\alpha_{it}^X}$$

- Markup $\mu_{it} \equiv \frac{P_{it}^Y}{\lambda_{it}}$,
- the output elasticity for X^v: $\theta_{it}^{X^v} \equiv \frac{\partial Q_{it}(.)}{\partial X_{it}} \frac{X_{it}}{Q_{it}}$
- X's expenditure share in total revenue $\alpha_{it}^{X^v} \equiv \frac{P^{X_{it}^v} X_{it}^v}{P_{it} Q_{it}}$.
  - Note that $\mu_{it} = \frac{\theta_{it}^{X^v}}{\alpha_{it}^X}$ holds for any variable input X_{it}.
  - In the majority of the empirical estimations, I use labor as the variable input. In that case, I define $\alpha_{it}^L \equiv \frac{VA_{it}}{w_t l_{it}}$, where VA_{it} is value added.
  - In some robustness checks, I employ materials as the variable input. In that case, I define $\alpha_{it}^M \equiv \frac{S_{it}}{p_t^M M_{it}}$, where S_{it} is sales and $p_t^M M_{it}$ is expenditure of materials.
- For Cobb-Douglas, θ_{it}^X is constant, so all within-sector variation is driven by α_{it}^X.

A.4 Further stylized Facts

Robustness on MRPK dispersion and productivity volatility

In this section, I follow Asker, Collard-Wexler, and De Loecker (2014) and implement their plant-level robustness check for examining the relationship between MRPK dispersion and productivity volatility. In general, the relationship here is in line with the findings in Asker, Collard-Wexler, and De Loecker (2014).

Table A.1: MRPK dispersion: plant-level robustness

	$MRPK_{irst}$ (Gross Revenue)								$MRPK_{irst}$ (Value Added)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
$a_{irst} - a_{irst-1}$	0.723**		0.666**		0.565**		0.441**		1.091**		1.053**		0.997**		0.935**	
	(0.0182)		(0.0149)		(0.00490)		(0.00408)		(0.00613)		(0.00300)		(0.00325)		(0.00332)	
$a_{irst-1} - a_{irst-2}$		0.527**		0.469**		0.239**		0.172**		0.612**		0.564**		0.250**		0.187**
		(0.0174)		(0.0163)		(0.00481)		(0.00480)		(0.0158)		(0.0153)		(0.00592)		(0.00601)
a_{irst-1}	0.906**		0.832**		0.660**		0.515**		1.136**		1.087**		1.017**		0.944**	
	(0.0156)		(0.0117)		(0.00533)		(0.00471)		(0.00826)		(0.00423)		(0.00380)		(0.00387)	
a_{irst-2}		0.756**		0.682**		0.309**		0.230**		0.863**		0.804**		0.325**		0.252**
		(0.0151)		(0.0139)		(0.00624)		(0.00603)		(0.0140)		(0.0139)		(0.00756)		(0.00764)
k_{irst}			-0.199**				-0.479**				-0.134**				-0.241**	
			(0.00491)				(0.00480)				(0.00621)				(0.00296)	
k_{irst-1}				-0.187**				-0.254**				-0.141**				-0.226**
				(0.00560)				(0.00843)				(0.00773)				(0.00906)
Plant FE	No	No	No	No	Yes	Yes	Yes	Yes	No	No	No	No	Yes	Yes	Yes	Yes
Observations	235765	161861	235765	161861	235765	161861	235765	161861	235765	147713	235765	147713	235765	147713	235765	147713

Standard errors in parentheses

SEs clustered at the sector-level for specifications 1-4, 9-12 and at the plant-level for specifications 5-8, 13-16.

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

Markups: urban and rural

Table A.2: Difference in markups: urban versus rural

	(1) ln(μ) - W	(2) ln(μ)- W	(3) ln(μ) - M	(4) ln(μ) - M
1(Urban)	-0.110***	-0.0367**	-0.00568	-0.0169***
	(0.0142)	(0.0115)	(0.00473)	(0.00479)
age		-0.0167***		0.00253***
		(0.000957)		(0.000225)
age^2		0.0000742***		-0.0000147***
		(0.0000163)		(0.00000285)
Constant	-2.094***	-1.884***	0.404***	0.376***
	(0.0178)	(0.0240)	(0.00541)	(0.00679)
Year FE	Yes	Yes	Yes	Yes
Observations	420422	398267	420755	398528

Standard errors in parentheses

SEs clustered at state-sector level. All specifications include Sector-state FEs

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A.5 Event Study on Dereservation: Urban/Rural Distinction

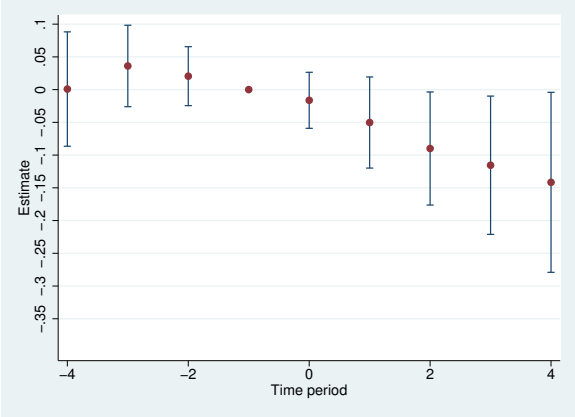
$$y_{irst} = \alpha_{rs} + \gamma_t + \zeta 1(Rural_{irs}) + \sum_{\tau=-4}^{4} \beta_\tau 1(t=\tau) + \sum_{\tau=-4}^{4} \beta_\tau^R 1(t=\tau) * 1(Rural_{irs}) + \varepsilon_{irst} \quad \text{(A.11)}$$

where $y_{irst} = \mu_{irst}, g(k_{irst})$ and where I bin up the end-points and normalize $\beta_{-1} = 0$. The reason why I investigate heterogeneity for rural plants, is that empirically, baseline markups are lower in an urban setting (see Table G.3). Therefore, an increase in competition might affect internally financed capital growth more for plants in an urban setting. In the empirical tests of the model predictions, the rural/urban distinction will be a relevant, though not essential, dimension of heterogeneity.[2]

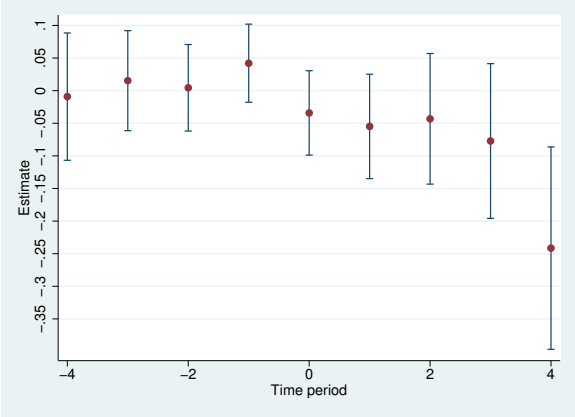
[2]Note that age and geographic location are almost the only contemporaneous dimensions of exogenous heterogeneity for incumbent plants. As such, analyzing heterogeneous treatment effects along this dimension is a valid empirical exercise.

Figure A.1: Dereservation Event-study on Markups and Capital Growth

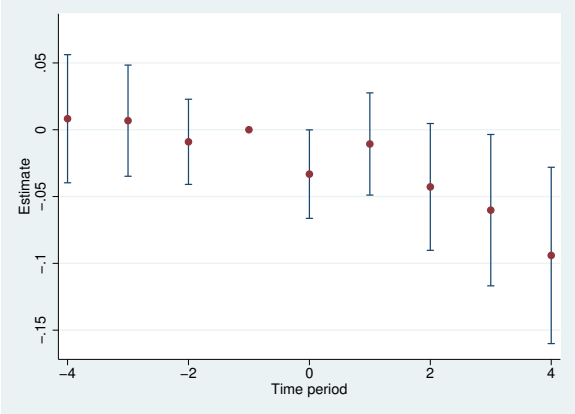
(a) Markup for Urban Plants



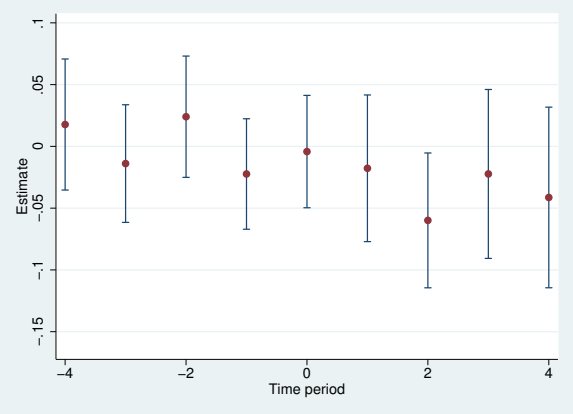
(b) Markup for Rural Plants



(c) $g(k_{irst})$ for Urban Plants



(d) $g(k_{irst})$ for Rural Plants



The figure displays the coefficients and 95% confidence intervals of an event-study regression on dereservation. Panels (a,b) display the results of the regression $\mu_{irst} = \alpha_{irs} + \gamma_t + \zeta 1(Rural_{irs}) + \sum_{\tau=-4}^{4} \beta_\tau 1(t = \tau) + \sum_{\tau=-4}^{4} \beta_\tau^R 1(t = \tau) * 1(Rural_{irs}) + \varepsilon_{irst}$, while panels (c,d) display the results from the following regression: $g(k_{irst}) = \alpha_{rs} + \gamma_t + \zeta 1(Rural_{irs}) + \sum_{\tau=-4}^{4} \beta_\tau 1(t = \tau) + \sum_{\tau=-4}^{4} \beta_\tau^R 1(t = \tau) * 1(Rural_{irs}) + \varepsilon_{irst}$. Panels (a,c) display the results for β_τ, where I normalize $\beta_{-1} = 0$. Panels (b,d) show estimates for β_τ^R.

A.6 Capital-labor ratio convergence

The main proposition in the theory section predicts that capital wedges shrink faster in a market with lower levels of competition. In this appendix section, I present additional evidence for this prediction, arising from the convergence of plant-level capital-labor ratios to their optimal level.

From the expressions in the theory section, combining the expression for optimal labor choice:

$$l_{it} = \mu_{it}^{\frac{1}{\eta-1}} \frac{Q}{M} a_{it}^{\frac{\eta}{1-\eta}} \left(\frac{P(1-\alpha)}{w} \right)^{\frac{1-\eta+(1-\alpha)\alpha\eta^2}{(1-\eta)(1+\alpha\eta-\eta)}} \left(\frac{\alpha}{\omega_{it}} \right)^{\frac{\alpha}{1-\eta}} \tag{A.12}$$

and equation (1.25) for optimal capital choice, one can find that the capital labor ratio takes the following form:

$$\frac{k_{it}}{l_{it}} = \frac{(1-\alpha)\alpha}{\omega_{it}} \frac{P}{w} \tag{A.13}$$

As such, theoretically the only source of variation in $\frac{k_{it}}{l_{it}}$ across firms within a sector arises from the capital wedges ω_{it}. In the table below I test whether the speed of convergence of the capital-labor ratio is faster in settings with less competition. The data again confirm this prediction of the model.

Table A.3: Capital-labor ratio: speed of convergence

	$\left(\frac{k}{l}\right)_{irst}$	
	(1)	(2)
$\left(\frac{k}{l}\right)_{irst-1}$	0.337**	0.315**
	(0.00472)	(0.00714)
$\left(\frac{k}{l}\right)_{irst-1}$ * $Median_{rst-1}[\ln \mu_{irst-1}]$	-0.0216**	-0.00221
	(0.00548)	(0.00885)
$\left(\frac{k}{l}\right)_{irst-1}$ * $Fin\ Dep_s$		-0.0109
		(0.0132)
$\left(\frac{k}{l}\right)_{irst-1}$ * $Median_{rst-1}[\ln \mu_{irst-1}]$ * $Fin\ Dep_s$		-0.0290
		(0.0162)
Influence of $Median_{rst-1}[\ln \mu_{irst-1}]$ on convergence speed:		
$\rho_1 * [90\%ile[Median(\ln \mu)] - 10\%ile[Median(\ln \mu)]]$	-0.0407	-0.0042
$[\rho_1 + \rho_3 * Fin\ Dep_s(90\%ile)] *$ $(90\%ile[Median(\ln \mu)] - 10\%ile[Median(\ln \mu)])$		-0.0461
Plant FE	Yes	Yes
State-sector-year FE	Yes	Yes
Observations	237344	193016

Standard errors, clustered at the plant-level, in parentheses (* $p < 0.05$, ** $p < 0.01$). The variable $\left(\frac{k}{l}\right)_{irst}$ is the firm-level capital-labor ratio, in logs. The inverse measure for competition, $Median_{rst}[\ln \mu_{irst}]$, is demeaned within sectors. Both specifications include a cubic polynomial in age as control variables.

$90\%ile[Median(\ln \mu)]$ and $10\%ile[Median(\ln \mu)]$ are the respective values for the 90th and the 10th percentile of $Median_{rst-1}[\ln \mu_{irst-1}]$ across state-sector-year observations. This way, $\rho_1 * [90\%ile[Median(\ln \mu)] - 10\%ile[Median(\ln \mu)]]$ reports the difference in average convergence rate for firms exposed to the value of the median markup in the respective percentiles. In specification (2), this is for firms in sectors with 0% financial dependence.

$90\%ile[Median(\ln \mu)] * [\rho_1 + \rho_3 * Fin\ Dep_s(90\%ile)] - 10\%ile[Median(\ln \mu)] * [\rho_1 + \rho_3 * Fin\ Dep_s(90\%ile)]$ reports the difference in average converge rates, due to different median markups, for firms producing in sectors at the 90th percentile of financial dependence

Appendix B

Slicing the Pie

B.1 Trade and Aggregate Sectoral Reallocation in Germany

In this appendix section, we provide descriptives on the changing composition of output across sectors and how these compositional changes are related to trade. Specifically, we decompose the changes in sectoral shares of total output into changes in domestic demand and changes in net exports. This descriptive exercise will demonstrate and visualize the substantial magnitude of sectoral reallocation, and at the same time quantify the relative importance of changes in net exports in this reallocation. We then examine how the observed changes in output shares relate to shifts in sectoral employment shares. Taken together, this exercise provides strongly suggestive evidence for trade-induced sectoral reallocation.

Decomposition of Sectoral Reallocation

We start from the accounting identity

$$E_{is}^t = Y_{is}^t - X_{is}^t + M_{is}^t,$$

where E_{is}^t is country $i's$ expenditure in sector s at time t, Y_{is}^t is production, X_{is}^t is exports and M_{is}^t is imports. Rearranging and dividing both sides by total expenditure in country i yields

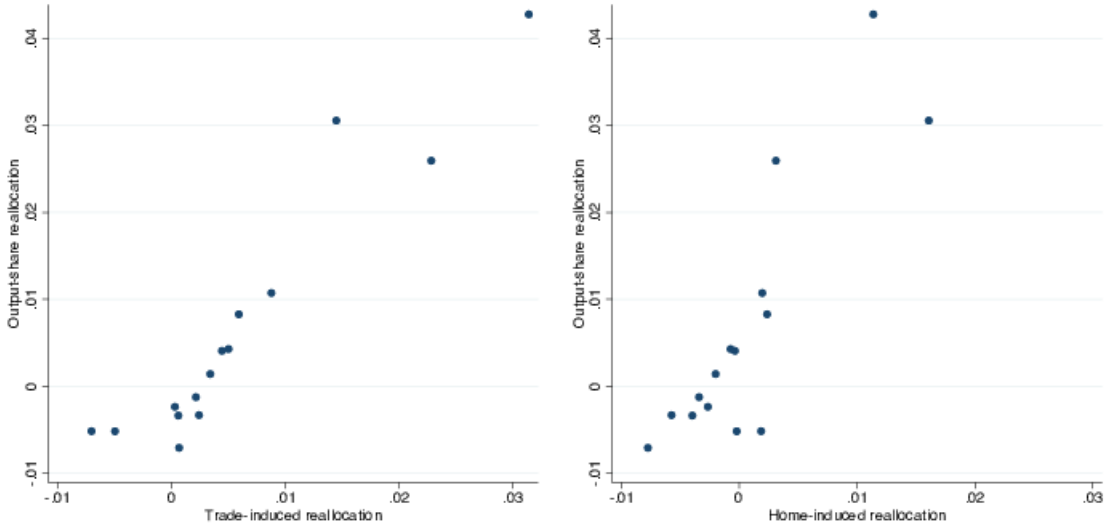
$$\frac{Y_{is}^t}{E_i^t} = \frac{E_{is}^t - M_{is}^t + X_{is}^t}{E_i^t} = \beta_{is}^t \lambda_{is}^t + \frac{X_{is}^t}{E_i^t},$$

where $\beta_{is}^t \equiv \frac{E_{is}^s}{E_i^t}$ are expenditure shares across goods and $\lambda_{is}^t \equiv \frac{E_{is}^t - M_{is}^t}{E_{is}^t}$ is the domestic trade share in sector s. Changes over time in $y_{is}^t \equiv \frac{Y_{is}^t}{E}$ can be decomposed as

$$\underbrace{y_{is}^t - y_{is}^{t-1}}_{\text{"Output-share" reallocation}} = \underbrace{(\beta_{is}^t - \beta_{is}^{t-1})\lambda_{is}^t}_{\text{"Home-related" reallocation}} + \underbrace{(\lambda_{is}^t - \lambda_{is}^{t-1})\beta_{is}^{t-1} + \frac{X_{is}^t}{E_i^t} - \frac{X_{is}^{t-1}}{E_i^{t-1}}}_{\text{"Trade-related" reallocation}}.$$

To bring this equation to the data, we focus on Germany and set $t = 2007, t-1 = 2000$. We first visualize the decomposition of changes in output shares in Figure **??**, for 15 manufacturing sectors at the 2-digit level of aggregation.[1] We see that both trade-related and home-related reallocation are strongly correlated with output-share reallocation. The sector with the highest output-share reallocation, with an increase of 3.9 percentage points, is the sector producing "Motor Vehicles, Trailers, and Semi-Trailers."

Figure B.1: Decomposition of Changes in Output Shares



We now quantify the share of trade-related and home-related reallocation in the ouput-share reallocation. Define $G_{is}^t \equiv y_{is}^t - y_{is}^{t-1}$, $H_{is}^t \equiv (\beta_s^t - \beta_s^{t-1})\lambda_{is}^t$, $T_{is}^t \equiv (\lambda_{is}^t - \lambda_{is}^{t-1})\beta_s^{t-1} + \frac{X_{is}^t}{C_i^t} - \frac{X_{is}^{t-1}}{C_i^{t-1}}$, such that $G_{is}^t = H_{is}^t + T_{is}^t$. We want to know what share of the variance of changes in output shares (G_{is}^t) is home-related (i.e. related to H_{is}^t), and what share is trade-related (i.e. related to T_{is}^t). We can answer this question by running two separate regressions where we either regress H_{is}^t on G_{is}^t, or T_{is}^t on G_{is}^t.[2] The results are shown in Table **??**. Around 64% of the variance of changes in output shares is due to changes in trade-related reallocation, while the remainder is related to home-related reallocation.

[1] Section 3.3 provides a detailed discussion of the data.

[2] Formally, we run the following regressions: $H_{is}^t = \alpha + \beta_1 G_{is}^t + \epsilon$; $T_{is}^t = \alpha + \beta_2 G_{is}^t + \epsilon$, so $\beta_1 = cov(G_{is}^t, H_{is}^t)/var(G_{is}^t)$, $\beta_2 = cov(G_{is}^t, T_{is}^t)/var(G_{is}^t)$ and $\beta_1 + \beta_2 = 1$.

Table B.1: Decomposition of Changes in Output Shares

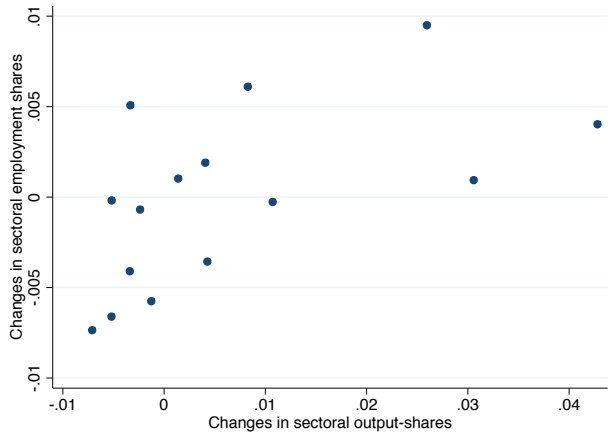
	(1) Trade-induced Reallocation	(2) Home-induced Reallocation
Output-share Reallocation	0.643***	0.357***
	(0.0583)	(0.0583)
Constant	0.00174	-0.00174
	(0.000927)	(0.000927)
Observations	15	15

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

As a final step, we ask to what extent changes in output shares, $y_{is}^t - y_{is}^{t-1}$, are correlated to changes in employment shares, $\pi_{is}^t - \pi_{is}^{t-1}$ with $\pi_{is}^t \equiv L_{is}^t / L_i^t$. Empirically, we find that there is a correlation of 56,8% between changes in sectoral output shares and changes in employment shares. We visualize this relation in Figure B.2.

Figure B.2: Relation between Sectoral Output and Employment Shares



B.2 US version

Data

For the US, we combine employment data from the County Business Patterns (CBP) dataset and sectoral output data from the NBER CES database. We also employ data on trade flows

and regional earnings that were kindly provided by Gordon Hanson.

We follow Autor, Dorn, and Hanson (2013) - (ADH) in defining regional economies using the concept of Commuting Zones (CZs). Our industry classification follows the 1987 SIC classification codes aggregated to the 2-digit level by an algorithm also provided by ADH, and restricted to manufacturing industries only. This leaves us with a total of 722 CZs and 20 industries. All current figures are for the year 2000.

For employment shares π_{igs}, we apply the same algorithm as ADH to obtain commuting zone employment shares from the CBP county level data. As in the German case, we currently input very low values ($\pi_{igs} = e^{-10}$) to CZ-industry cells with zero values. Our figures for national sectoral output Y_{is} come directly from the NBER-CES database variable *vship*, which represents the total value of industry shipments. To obtain aggregate earnings in manufacturing at the CZ level (Y_{ig}), we employ publicly available data from ADH's China Syndrome paper. Specifically, we multiply each commuting zone's weekly average wages in manufacturing by their employment count in manufacturing.

Decomposition

Here, we implement the same analysis as in Section 2.1, but now for the US, with $t = 2007, t-1 = 1995$.

Figure B.3: Decomposition of Changes in Output Shares - US

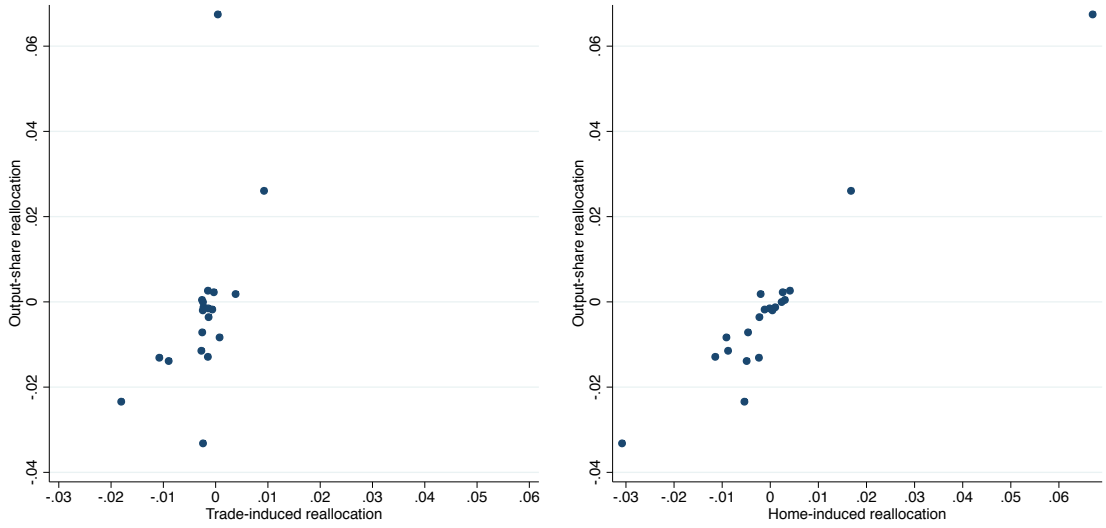


Figure A.1. displays the relation between output-share reallocation and first trade-induced reallocation (Left Panel) and second home-induced reallocation (Right Panel.) One sector is an important outlier in terms of output-share reallocation, namely "Petroleum Refining and Related Industries." Since this outlier is largely explained by home-induced

reallocation, trade-induced reallocation will play a smaller role in the US. This is demonstrated by the regression results in Table A.1., where only 13.6% of output-share reallocation is trade-induced, substantially below the 64.3% in Germany.

Table B.2: Decomposition of Changes in Output Shares

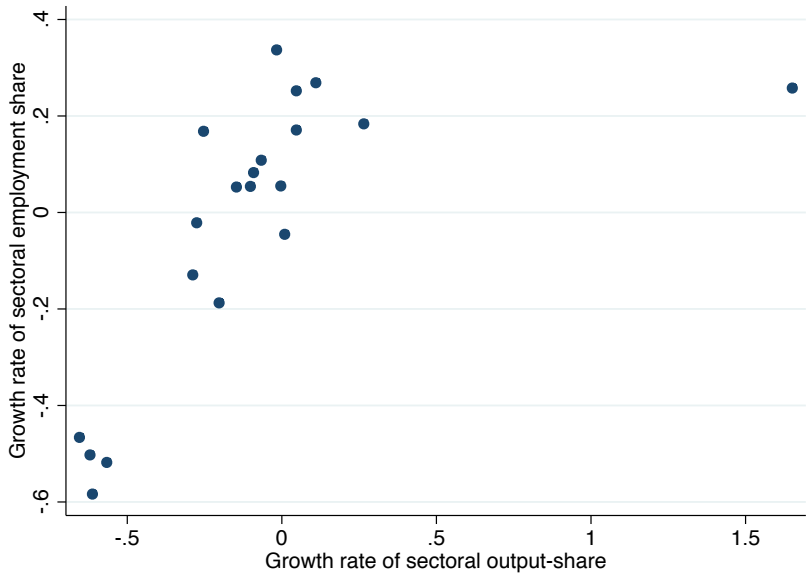
	(1) Trade-induced Reallocation	(2) Home-induced Reallocation
Output-share Reallocation	0.136*	0.864***
	(0.0553)	(0.0553)
Constant	-0.00216	0.00216
	(0.00109)	(0.00109)
Observations	20	20

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure A.2. shows the relation between growth rates of sectoral output-shares and growth rates of sectoral employment shares. The correlation of these growth rates is 66.5% in our US data.

Figure B.4: Relation between Sectoral Output and Employment Shares



Autarky Exercise - US

Figure B.5: Distributional Gains by Region - Autarky - US

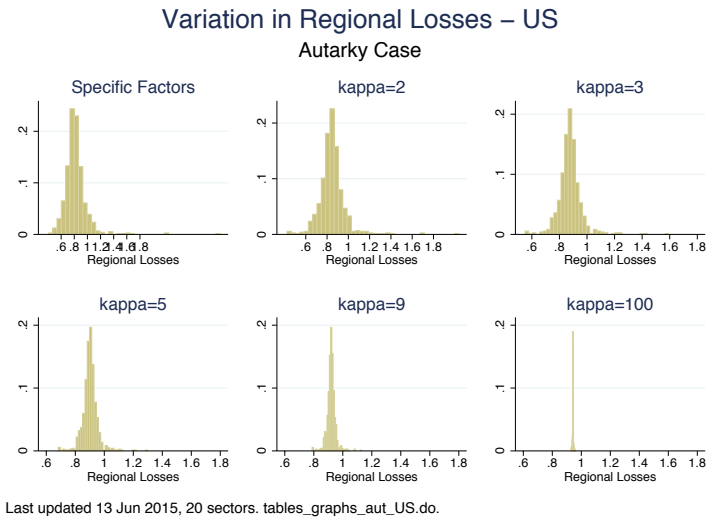


Figure B.6: Distributional Gains by Region - Autarky - US

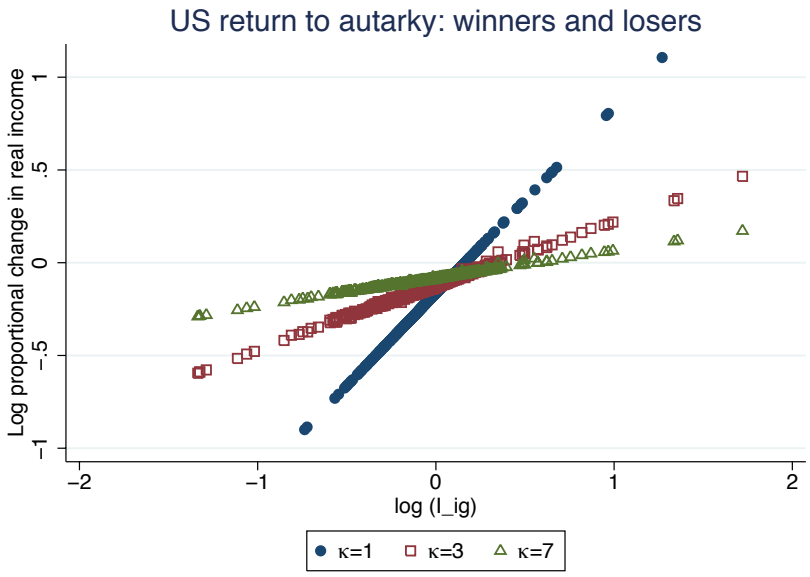
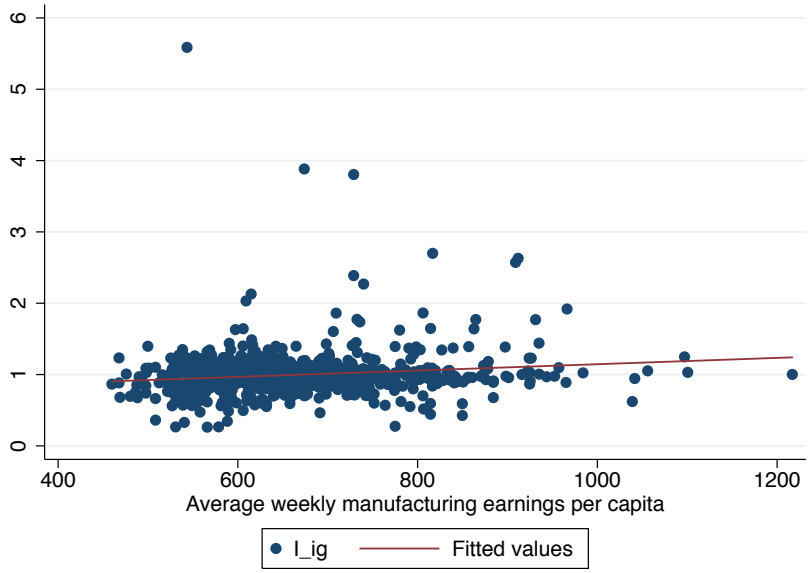


Figure B.7 provides insight into why the inequality-adjusted gains from trade are strongly positive in the US. First, the correlation between import-competition and income per capita is positive, at 0.147. Hence, on average poorer regions gain more from trade than richer regions,

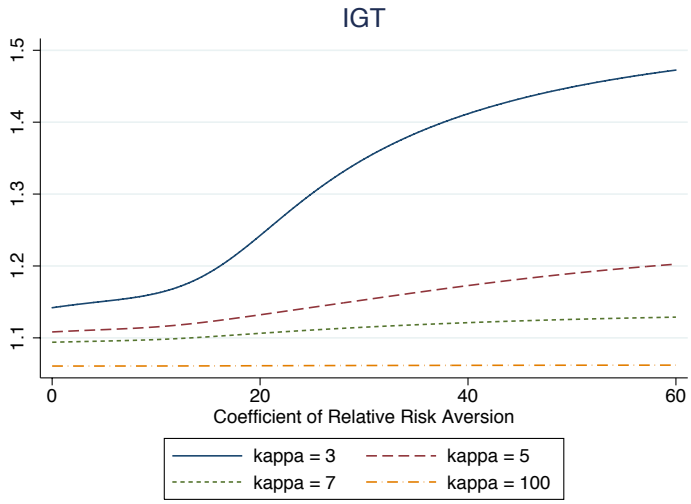
such that trade is pro-poor. In these bottom percentiles, the export-oriented regions are well represented. These export-oriented regions unambiguously lose from going to autarky, whereas the import-competing regions lose less. As such, inequality among the bottom income percentiles is mitigated under trade, compared to autarky.[3]

Figure B.7: Correlation between import-competition and earnings per worker



[3]We are in the process of exploring the robustness of these result to model-specifications that include capital in the production function or for differences in skills across worker groups and to different measurements of import-competition.

Figure B.8: IGT - US



B.3 First-stages of estimation procedures

Figure B.9: First stage for Table 3.4

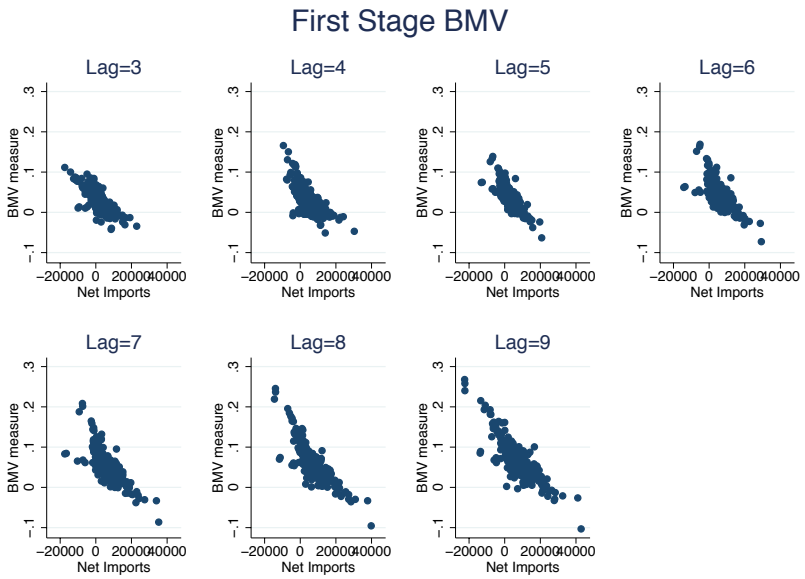
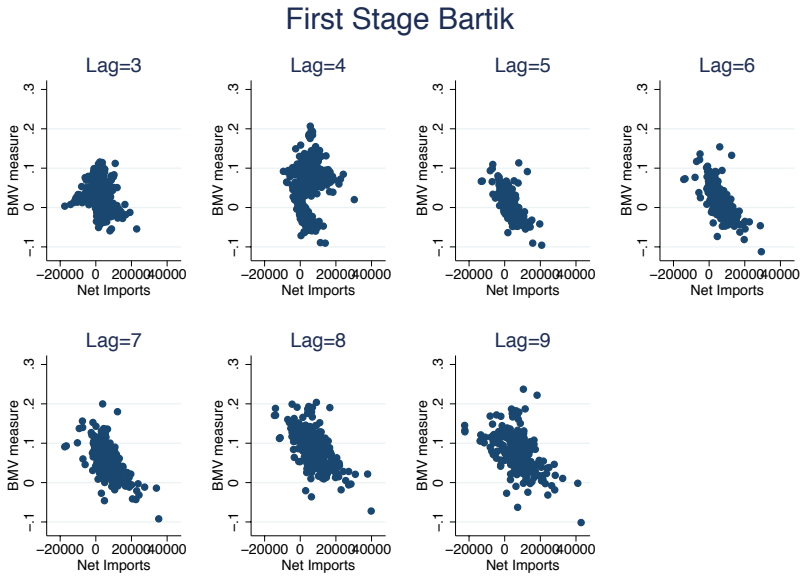


Figure B.10: First stage for Table 3.3



B.4 Extensions with home production

Regional income and reallocation

Here, we provide an extension of our estimation procedure in Section 3.4, which includes a non-tradable sector.

Derivation

Start from

$$Y_{igM} \equiv \gamma \frac{\sum_{k=1}^{S} A_{igk} w_{ik}^{\kappa}}{\Phi_{ig}^{\kappa}} \Phi_{ig} L_{ig} = \gamma \pi_{igM} \Phi_{ig} L_{ig}$$

with $\Phi_{ig}^{\kappa} = \sum_{k=0}^{S} A_{igk} w_{ik}^{\kappa}, \pi_{igM} = \frac{\sum_{k=1}^{S} A_{igk} w_{ik}^{\kappa}}{\Phi_{ig}^{\kappa}}$. Therefore:

$$\frac{Y_{igM}}{\pi_{igM} L_{ig}} = \gamma \Phi_{ig}$$

Here, define $y_{igM} \equiv \frac{Y_{igM}}{\pi_{igM} L_{ig}}$ as the average income per manufacturing worker. We then have that:

$$\hat{y}_{igM} = \hat{\Phi}_{ig}$$

Taking logs:

$$\ln \hat{y}_{igM} = \ln \hat{\Phi}_{ig} \tag{B.1}$$

This equation is very closely related to our previous approach, except that we now need to update the derivation of the observable counterpart to $\ln \hat{\Phi}_{ig}$. To this end, note that for $s \geq 1$ we have $\hat{\pi}_{igs} = \hat{w}_{is}^{\kappa} / \hat{\Phi}_{ig}^{\kappa}$ and $\hat{\pi}_{ig0} = w_{ig0}^{\kappa} / \hat{\Phi}_{ig}^{\kappa}$ and hence

$$\begin{aligned} \frac{\hat{w}_{is}^{\kappa}}{\hat{w}_{i1}^{\kappa}} &= \frac{\hat{\pi}_{igs}}{\hat{\pi}_{ig1}} \\ \frac{\hat{w}_{ig0}^{\kappa}}{\hat{w}_{i1}^{\kappa}} &= \frac{\hat{\pi}_{ig0}}{\hat{\pi}_{ig1}} \end{aligned}$$

Combining these expressions we get

$$\frac{1}{\kappa} \ln \left(\sum_{s} \pi_{igs} \frac{\hat{\pi}_{igs}}{\hat{\pi}_{ig1}} \hat{w}_{i1}^{\kappa} \right) = \frac{1}{\kappa} \ln \left(\frac{\hat{w}_{i1}^{\kappa}}{\hat{\pi}_{ig1}} \right)$$

The problem is that $\frac{\hat{w}_{is}^{\kappa}}{\hat{w}_{i1}^{\kappa}} = \frac{\hat{\pi}_{igs}}{\hat{\pi}_{ig1}}$ would make this very sensitive to the choice of g . We can use the fact that for $s \geq 1$ this relation $\left(\frac{\hat{w}_{is}^{\kappa}}{\hat{w}_{i1}^{\kappa}} = \frac{\hat{\pi}_{igs}}{\hat{\pi}_{ig1}} \right)$ holds for all g, so that for $s \geq 1$ we have

$$\frac{\hat{w}_{is}^{\kappa}}{\hat{w}_{i1}^{\kappa}} = \nu_{is}(1)$$

where for any $k \geq 1$ we have

$$\nu_{is}(k) \equiv \exp \frac{1}{G} \sum_{g} \log \frac{\hat{\pi}_{igs}}{\hat{\pi}_{igk}}$$

But note that $\frac{\hat{w}^{\kappa}_{ig0}}{\hat{w}^{\kappa}_{i1}} = \frac{\hat{\pi}_{ig0}}{\hat{\pi}_{ig1}}$ does not mean that $\frac{\hat{w}^{\kappa}_{ig0}}{\hat{w}^{\kappa}_{i1}} = \frac{\hat{\pi}_{ih0}}{\hat{\pi}_{ih1}}$ for $h \neq g$, so we cannot do this for all s and say that

$$\sum_{k=0}^{S} \pi_{igk} \hat{w}^{\kappa}_{igk} = \hat{w}^{\kappa}_{i1} \sum_{k=0}^{S} \pi_{igk} v_{is}(1)$$

Instead we need to use:

$$\pi_{ig0} \hat{w}^{\kappa}_{ig0} + \sum_{s \geq 1} \pi_{igs} \hat{w}^{\kappa}_{is} = \hat{w}^{\kappa}_{i1} \left[\pi_{ig0} \frac{\hat{\pi}_{ig0}}{\hat{\pi}_{ig1}} + \sum_{s \geq 1} \pi_{igs} \nu_{is}(1) \right]$$

In order to reduce sensitivity to the reference sector 1, we can use that we have for all $s, k \geq 1$ that

$$\hat{w}^{\kappa}_{is} = \hat{w}^{\kappa}_{ik} \nu_{is}(k)$$

This implies that

$$\hat{w}^{\kappa}_{is} = \left(\exp \frac{1}{S} \sum_{k \geq 1} \log \hat{w}^{\kappa}_{ik} \right) \nu_{is}$$

where

$$\nu_{is} \equiv \exp \left(\frac{1}{S} \sum_{k \geq 1} \log \nu_{is}(k) \right)$$

We now have

$$\pi_{ig0} \hat{w}^{\kappa}_{ig0} + \sum_{s \geq 1} \pi_{igs} \hat{w}^{\kappa}_{is} = \pi_{ig0} \hat{w}^{\kappa}_{ig0} + \left(\exp \frac{1}{S} \sum_{k \geq 1} \log \hat{w}^{\kappa}_{ik} \right) \sum_{s \geq 1} \pi_{igs} \nu_{is}$$

Using the fact that for any $k \geq 1$ we have

$$\hat{w}^{\kappa}_{ig0} = \frac{\hat{\pi}_{ig0}}{\hat{\pi}_{igk}} \hat{w}^{\kappa}_{ik}$$

we can then write

$$\begin{aligned} \hat{w}^{\kappa}_{ig0} &= \exp \log \prod_{k \geq 1} \left(\frac{\hat{\pi}_{ig0}}{\hat{\pi}_{igk}} \right)^{1/S} \prod_{k \geq 1} (\hat{w}^{\kappa}_{ik})^{1/S} \\ &= \left(\exp \frac{1}{S} \sum_{k \geq 1} \log \hat{w}^{\kappa}_{ik} \right) \varphi_{ig} \end{aligned}$$

where

$$\varphi_{ig} \equiv \left(\exp \frac{1}{S} \sum_{k \geq 1} \log \left(\frac{\hat{\pi}_{ig0}}{\hat{\pi}_{igk}} \right) \right)$$

so finally we have

$$\pi_{ig0} \hat{w}_{ig0}^{\kappa} + \sum_{s \geq 1} \pi_{igs} \hat{w}_{is}^{\kappa} = \left(\exp \frac{1}{S} \sum_{k \geq 1} \log \hat{w}_{ik}^{\kappa} \right) \left(\pi_{ig0} \varphi_{ig} + \sum_{s \geq 1} \pi_{igs} \nu_{is} \right) \tag{B.2}$$

Estimating equation and instruments

We obtain our estimating equation by substituting equation (B.2) into equation (B.1):

$$\ln \hat{y}_{igM} = b_i + \frac{1}{\kappa} \ln \left(\pi_{ig0} \varphi_{ig} + \sum_{s \geq 1} \pi_{igs} \nu_{is} \right) + \varepsilon_{ig} \tag{B.3}$$

Here, as an instrument for $\ln \left(\pi_{ig0} \varphi_{ig} + \sum_{s \geq 1} \pi_{igs} \nu_{is} \right)$ we can use the usual Bartik instrument $\sum_s \pi_{igs} \frac{\Delta IP_{st}^{East \to Other}}{L_{is}}$.