Essays in International Economics

by

Ting Lan

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Economics)
in The University of Michigan
2020

Doctoral Committee:

Assistant Professor Dominick Bartelme, Co-Chair
Professor Andrei Levchenko, Co-Chair
Assistant Professor Javier Cravino
Professor Jagadeesh Sivadasan

Ting Lan
tinglan@umich.edu
ORCID iD: 0000-0002-2329-0632

To Mom and Dad, who are the first and best teachers in my life.

Acknowledgments

I am grateful to many people, who have provided tremendous help along the way during my Ph.D. journey. Without their support, I could not archive any accomplishment in my Ph.D. life.

I am deeply indebted to my dissertation committee - Andrei Levchenko, Dominick Bartelme, Javier Cravino, and Jagadeesh Sivadasan - for their invaluable guidance, unconditional encouragement, and endless patience. Andrei Levchenko has spent countless hours advising all my research projects and provided many helpful suggestions and solutions to my research problems in various stages. Dominick Bartelme helped me immensely navigate the trade and economic geography literature and inspired me to come up with the original idea for my job market paper. Javier Cravino helped me reshaped my job market paper and spotted the most important questions that my paper is suitable to answer. Jagadeesh Sivadasan taught me structural modeling, which opened another door of research opportunities and added an invaluable part to my dissertation. All of them have been my role models as serious thinkers and passionate researchers.

I am additionally grateful for Sebastian Stole, Ying Fan, and David Miller, who have provided exceptional support largely above their duties. I have also hugely benefited from discussing my dissertation with Ruzic Dimitrije, Laurien Gilbert, Xing Guo, and Nicolas Morales.

I have learned much from my fantastic colleagues at Michigan: Luis Baldomero-Quintana, Hannah Bolder, Thomas Brosy, Kimberly Conlon, Yingjie Feng, Tangren Feng, Chiara Ferrero, Bhanu Gupta, Anirudh Jayanti, Xinwei Ma, Alexandr Moskalev, Merve Sariisik, Shuqiao Sun, Andrew Usher, Robert Venyige, Pinghui Wu, Huayu Xu, Wenjian Xu, Hang Yu.

I very much appreciate the generous financial support from the Department of Economics, and the great administrative assistance from Hiba Baghdadi, Laura Flak, Julie Heintz, and Cassandra Kelly.

Finally and most importantly I want to extend my deepest gratitude to my family. My parents, who encouraged me to apply for the Ph.D. programs in the US, and

always has full faith in me. My husband, Dayin Zhang, for his unconditional love, support, and understanding over the past six years. My American Mom, Teresa Lynn, who helped me get in and get out of graduate school.

Table of Contents

List of Tables

Table

List of Figures

Figure

List of Appendices

Appendix

Abstract

This dissertation lies at the intersection of international trade and macroeconomics, and it related to topics on economic growth, innovation and productivity, monetary policy, and economic geography.

Chapter 1 evaluates and quantifies the importance of production proximity for innovation efficiency. First, using a novel and comprehensive establishment-level dataset from the U.S. Census Bureau, I document that innovation and production activities are coagglomerated in the majority of manufacturing industries. The geographic concentration of innovation and production provides suggestive evidence on the importance of the two activities being colocated. Second, I develop an empirical model that allows for spillovers from local production to innovation and apply it to measure the private returns to R&D investment for a panel of U.S. R&D firms during 2002-2012. My estimates show that the proximity to production raises the returns to R&D, suggesting there are positive spillovers from the local manufacturing to innovation. Third, I evaluate the macroeconomic implications of my empirical findings in a multi-region production and trade model featuring the local spillovers from production to innovation. I find that the relocation of production workers due to the China trade shock leads to a moderate reduction in both process and product innovation. States with a larger decline in manufacturing employment experience a more substantial loss in innovation efficiency.

Chapter 2 studies differential responses of prices faced by different consumers following macroeconomic shocks. Monetary shocks have distributional consequences if they affect relative prices across goods consumed by different households. We document that the prices of the goods consumed by high-income households are stickier and less volatile than those of the goods consumed by middle-income households. Following a monetary policy shock, the estimated impulse responses of high-income households' consumer price indices are about one-third smaller than those of the middle-income households. We evaluate the implications of these findings in a quantitative multi-sector New-Keynesian model featuring heterogeneous households. The distributional consequences of monetary policy shocks are large and similar to those

in the econometric model.

Chapter 3 estimates the impact of foreign sectoral demand and supply shocks on real income. Our empirical strategy is based on a first-order approximation to a wide class of small open economy models that feature sector-level gravity in trade flows. The framework allows us to measure foreign shocks and characterize their impact on income in terms of reduced-form elasticities. We use machine learning techniques to group 4-digit manufacturing sectors into a smaller number of clusters and show that the cluster-level elasticities of income with respect to foreign shocks can be estimated using high-dimensional statistical techniques. We find clear evidence of heterogeneity in the income elasticities of different foreign shocks. Foreign demand shocks in complex intermediate and capital goods have large impacts on real income, and both supply and demand shocks in capital goods have particularly large impacts in poor countries. Counterfactual exercises show that both comparative advantage and geography play a quantitatively large role in how foreign shocks affect real income.

Chapter 1

The Coagglomeration of Innovation and Production

1.1 Introduction

What factors can enhance innovation is a question of enormous interest, as creation of new technology is the key driver of long-run economic growth. There is a growing recognition that the transmission of knowledge plays an essential role in the development of new ideas (Cohen and Levinthal, 1989; Aghion and Jaravel, 2015). Going back to Marshall (1920), economists have argued that the transmission of knowledge is facilitated by geographic proximity. Geographic proximity not only improves the information flows across innovation activities, but also fosters the communications between the production workers and the researchers. While studies have focused on knowledge spillovers across innovation activities (Jaffe et al., 1993; Caragliu and Nijkamp, 2015), little attention has been paid to the potential spillovers from production to innovation. Understanding the role of production in fostering innovation is particularly relevant as the employment in US manufacturing has shrunk by nearly 30 percent since the year 2000. In this paper, I evaluate and quantify the importance of production proximity for innovation efficiency.

First, I document the geographic concentration of innovation and production activities for US manufacturing industries by using a novel and comprehensive plant-level data from the Census Bureau. To identify the innovation and production activities at the micro-level, I link the Longitudinal Business Database (LBD) with the Business R&D and Innovation Survey (BRDIS). It enables me to measure the spatial distribution of these activities.

I quantify the concentration pattern of innovation and production activities from two aspects. I begin by using the Ellison and Glaeser (1997, hereafter EG) metric of agglomeration to measure the location pattern of innovation and production activities, respectively. Then I use the EG metric of coagglomeration to measure how colocated the innovation and production activities are within an industry. I find that (i) innovation is more agglomerated than production in the majority of industries; and ii) innovation and production activities are coagglomerated both in the absolute sense and relative to the coagglomeration of industrial production. The spatial distribution of innovation and production suggests that gains from geographic concentration are more significant for innovation than production (Capello and Lenzi, 2014; Buzard et al., 2015). More importantly, it highlights the importance of locating innovation and production facilities close to each other.

Second, motivated by the stylized facts, I develop and estimate an empirical model of production and innovation within a manufacturing firm to quantify the private returns to R&D, measured in terms of productivity gains generated by a marginal increase in R&D. The primary goal is to assess if the returns to R&D are higher when innovation plants are in regions with more production workers from their own industry. Building on the recent work by Aw et al. (2011), Doraszelski and Jaumandreu (2013) and Bøler et al. (2015), the model assumes that a plant's revenue is subject to the plant-specific performance which evolves according to a Markov process. The increment in plant performance depends on the R&D investment of the plant itself, the interaction between its R&D investment and the local manufacturing employment, as well as the transfer of technology from the other R&D plants within the same firm. A unique feature of my model is to consider spillovers from the local manufacturing to innovation explicitly. The inclusion of the spillovers allows for the possibility of learning from the production process. Local production workers are a source of knowledge that can enhance plants' R&D efficiency, which, in turn, has an impact on plant performance.

I construct a unique plant-level panel data on R&D investment and domestic production for US R&D firms in the manufacturing sector from 2002-2012 for my estimation. These data allow me to observe the input and output for each production plant, and the R&D expenditures for each innovation plant.

My empirical results show that innovating plants obtain significantly higher returns to R&D if they are located in counties with more of their own industry production workers. It suggests that there are positive spillovers from the local production to innovation. All else equal, doubling the local own industry manufacturing employment increases the impact of a plant's own investment on its productivity by 21.4%. My analysis is informative for understanding the role of local manufacturing in enhancing the efficiency of innovation plants. It supports the view put forth in Naghavi and Ottaviano (2009) that feedback from manufacturing plants is important for research labs.

Third, I evaluate the macroeconomic implications of my empirical findings by extending the multi-region production and trade model developed by Arkolakis et al. (2018) with two key modifications: (i) allowing firms to increase productivity through R&D; and (ii) incorporating the spillovers from local manufacturing to innovation. Guided by my empirical findings, my model assumes that regions' capability in fostering innovation increases with their employment of production workers. Firms born in regions that are more capable of fostering innovation enjoy a higher return to R&D and spend more on innovation. New technologies created through R&D can be used in multi-region production (MP). Firms face a tradeoff between market proximity and production capability when choosing where to locate their production. Given the difference in regions' capability in innovation and production, the availability of MP leads some regions to specialize in production and others in innovation.

I take the model to the data in year 2012, and calibrate it to 48 states in the US and the rest of the world (ROW). My quantitative analysis uses the China shock to evaluate the effects of production reallocation on the innovation efficiency. I model the rise of China as the productivity shocks to the ROW, and use the predicted changes in the US imports from China during 1997-2012 to quantify the size of these productivity shocks. I find that the relocation of production workers due to the China trade shock leads to a moderate reduction in both process and product innovation. States with a larger decline in manufacturing employment experience a more substantial loss in innovation efficiency.

My analysis brings together three strands of the literature. First, it contributes to the literature on the agglomeration and coagglomeration of economic activities. Ellison and Glaeser (1997) and Duranton and Overman (2008) find that industrial production is geographically concentrated. The agglomeration of industrial production can be explained by the Marshall forces of labor pooling, input sharing and knowledge spillovers (Ellison et al., 2010; Faggio et al., 2017). Much less is known about the coagglomeration of innovation and production activities due to limitation in data availability (Audretsch and Feldman, 1996; Carlino and Kerr, 2015). I contribute to this line of research by using a novel micro-level dataset to uncover the spatial distribution of innovation and production activities.

Second, my analysis is closely related to the work on R&D investment and plant productivity. Building on the knowledge capital model by Griliches et al. (1979), as well as more recent work by Aw et al. (2011), Doraszelski and Jaumandreu (2013), and Bøler et al. (2015), I evaluate the impact of R&D investment on plant productivity.

The focus on multi-unit firms also relates my analysis to Bilir and Morales (2016), which allows for the intra-firm transfer of technology. My approach is novel in its explicit consideration of the local employment impact on innovation efficiency. The estimates support the theories proposed by Duranton and Puga (2001) and Naghavi and Ottaviano (2009) that production plays an important role in fostering innovation.

Third, my paper also contributes to the studies that seek to understand the consequences of trade shocks for innovation. The empirical evidence about the impact of trade shocks on innovation is mixed. While Bloom et al. (2016) finds that European firms facing higher levels of Chinese import competition create more patents, raise their IT intensity and increase their productivity, Dorn et al. (2016) shows that the foreign import competition reduces US patent production. I propose a new channel to evaluate the impact of trade shocks on innovation. Trade shocks affect local innovation through production reallocation. Innovation efficiency is enhanced by local manufacturing. The increased exposure to import competition leads to the decline of manufacturing, and thus reduces the local innovation efficiency. To quantify the impact of trade shocks on innovation through the new channel, I extend Arkolakis et al. (2018) by incorporating the local spillovers from innovation and production.

The rest of this paper is organized as follows. Section 1.2 documents stylized facts about the geographic concentration of innovation and production activities. Section 1.3 develops and estimates the empirical model to assess the spillovers from the local manufacturing to innovation. Section 1.4 extends the Arkolakis et al. (2018) model to evaluate the macroeconomic implications of my empirical findings. Section 1.5 concludes.

1.2 The Location Pattern of Innovation and Production

Economic activities are geographically concentrated to utilize the advantages of proximity.[1] Activities with more substantial gains from proximity tend to be more closely located (Ellison and Glaeser, 1997; Ellison et al., 2010). Thus, the concentration patterns of economic activities provide suggestive evidence on the importance of geographic proximity to these activities.

In this section, I use a novel plant-level dataset of innovation and production to

[1]According to Marshall (1920), economic activities are geographically concentrated to reduce the costs of obtaining inputs and supplying outputs, to share a broader labor market, and to enjoy intellectual or technology spillovers.

establish two important stylized facts of their agglomeration and coagglomeration patterns.

1.2.1 EG Metrics and Data

Measuring the spatial distribution of innovation and production activities has long been recognized as extremely difficult due to the lack of micro-level data (Audretsch and Feldman, 1996; Carlino and Kerr, 2015). In this section, I exploit detailed and comprehensive establishment-level data from two Census Bureau surveys —the Longitudinal Business Database (LBD) and the Business R&D and Innovation Survey (BRDIS)—to document the geographic concentration of innovation and production activities for manufacturing industries. The LBD covers the universe of establishments in the US and contains annual data on their employment, industry and geographic location. The BRDIS is a confidential firm-level survey conducted annually by the US Census Bureau in partnership with the National Science Foundation (NSF). It collects detailed information on firms' R&D activities including R&D-related employment, R&D expenditure, and the geographic location of domestic and foreign R&D performance. Linking the LBD with the BRDIS allows me to identify the innovation and production activities at the establishment level. Thus, it enables me to document the geographic concentration of innovation and production activities for manufacturing industries.

I measure the industry-level geographic concentration of innovation and production activities by using Ellison and Glaeser (1997) metrics.[2] The EG metrics are derived from a sequential profit-maximizing plant location choice model. It compares the degree of spatial concentration of economic activities in an industry with what would arise if these activities were randomly distributed across locations. I will focus on the four-digit manufacturing industries of the 2002 North American Industry Classification System (NAICS) and measure the concentration of these activities at the county level for the sample period 2008-2012.

I quantify the concentration pattern of innovation and production activities in two ways. First, I use the EG metric of agglomeration to measure the location pattern of innovation and production activities, respectively. The EG agglomeration measure for

[2]Another widely used coagglomeration metric is the continuous measure developed by Duranton and Overman (2005).

economic activity A in industry k is

$$\gamma_k^A \equiv \frac{\sum_l \left(s_{lk}^A - x_l\right)^2 - \left(1 - \sum_l x_l^2\right) H_k^A}{\left(1 - \sum_l x_l^2\right)\left(1 - H_k^A\right)} \tag{1.1}$$

where $A \in \{P, I\}$ denotes the type of economics activities. $A = P$ when measuring the agglomeration pattern for production activities and $A = I$ for innovation ones. s_{lk}^A is the production (innovation) employment share in industry k at county l and x_l is county l's share of total population.[3] H_k^A is the Herfindahl index of the industry k's production (innovation) establishment size distribution.[4]

The EG metric of agglomeration measures the tendency for production (innovation) activities to be closely located. $\gamma_k^A = 0$ is a no-agglomeration benchmark such that production (innovation) activities are randomly located given the profitability of each county l. $\gamma_k^A = 1$ indicates that production (innovation) activities are perfectly agglomerated with all the production (innovation) employment in a single county.

Second, I use the EG metric of coagglomeration to quantify how collocated the innovation and production activities are within an industry.[5] The EG coagglomeration measure of innovation and production for industry k takes the form:

$$\gamma_k^c \equiv \frac{\sum_l \left(s_{lk}^P - x_l\right)\left(s_{lk}^I - x_l\right)}{1 - \sum_l x_l^2} \tag{1.2}$$

where s_{lk}^P is the production employment share in industry k at county l, s_{lk}^I is the innovation employment share in industry k, and x_l is the population share of county l.

The EG coagglomeration measure captures the tendency for innovation activities to locate near production ones. It is closely related to the covariance of the county's employment shares in innovation and production. Negative values of the coagglomeration measure arise when innovation and production activities are agglomerated in different areas. The coagglomeration measure is zero when the production and production activities are randomly located.

[3]In Ellison and Glaeser (1997),$x_l = \frac{\bar{\pi}_l}{\sum_l \bar{\pi}}$, where π_l is a random variable reflecting the profitability of locating in county l. In practice, x_l could be measured as the county l's share of overall manufacturing employment or county l's share of the total population. As long as, a county with a higher level of x_l will have a higher level of profits.

[4]The subtraction of H_k^A is an adjustment that accounts for the fact that $\sum_l \left(s_{lk}^A - x_l\right)^2$ measure is expected to be larger in industries consisting of fewer larger plants if locations were chosen completely at random.

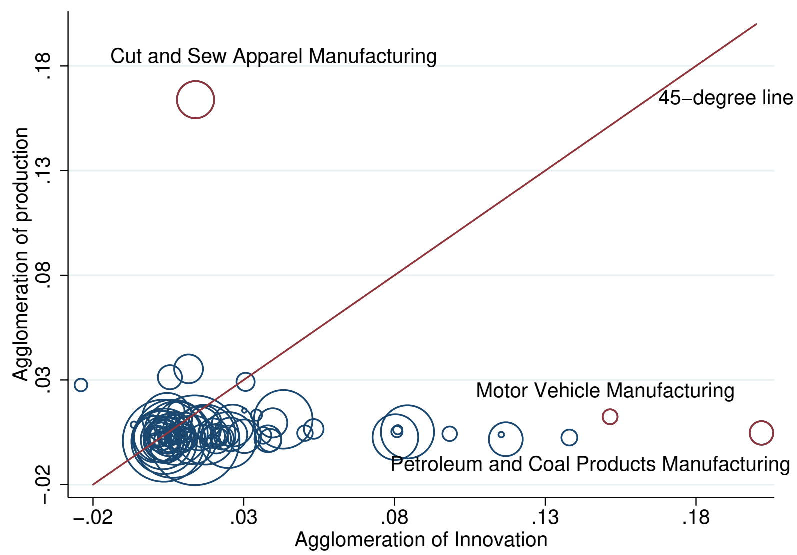
[5]The EG coagglomeration metric takes a simpler form when applied to measure the concentration of two activities. See appendix for the relationships between EG coagglomeration measure for two economic activities and for a group of activities.

1.2.2 Two Facts on the Spatial Distribution of Innovation and Production

Fact 1: Innovation activities are more agglomerated than production ones in the majority of manufacturing industries.

Figure 1.1 plots the agglomeration of production in each manufacturing industry against that of its innovation. Each circle represents a four-digit NAICS industry, and the size of the circle reflects the size of the industry. The x-axis measures the agglomeration pattern of innovation activities, and the y-axis measures that of production ones. The solid line is the 45-degree line. Most of the circles lie under the 45-degree line, implying that innovation activities are geographically more concentrated than production ones in the majority of manufacturing industries. This pattern suggests that the gains from concentration are more significant for innovation than production activities.

Figure 1.1 Agglomeration Indices of Innovation and Production

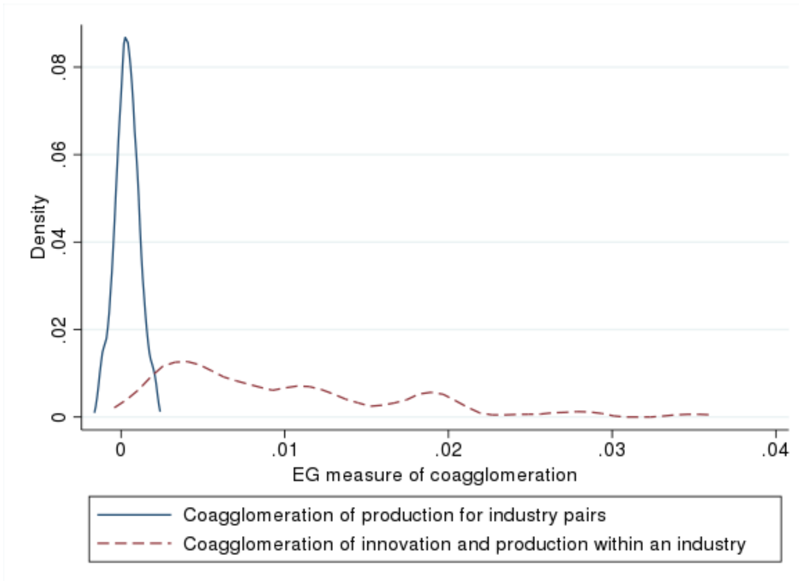


Notes: The figure plots the agglomeration indices of production in each manufacturing industry against agglomeration indices of innovation. Each circle represents a four-digit NAICS manufacturing industry, and the size of the circle reflects the size of the industry. The solid line is the 45-degree line.

Fact 2: Innovation and production activities are coagglomerated in the majority of manufacturing industries.

Figure 1.2 plots the density of the coagglomeration measures. The dash line plots the distribution of the coagglomeration measures for innovation and production. The majority of coagglomeration measures is greater than 0, implying that innovation

Figure 1.2 Kernel Density Estimates of the Coagglomeration Indices



Notes: The figure plots the Kernel density estimates of the coagglomration indices. The dash line plots the distribution of the coagglomeration measure for innovation and production. The solid line reports the distribution of the pairwise cross-industry coagglomeration measures for production.

tends to locate closely with production in most manufacturing industries.

To evaluate the strength of the coagglomeration between innovation and production within each industry, I compare it with the coagglomeration measures of industrial production. Production activities are known to be closely colocated across industries (Ellison et al., 2010; Duranton and Overman, 2008; Faggio et al., 2017)[6], and therefore, the coagglomeration of production serves as a good benchmark. For each industry, I compute the pairwise coagglomeration measures for its production with other manufacturing industries' production. [7] The solid line reports the distribution of the pairwise cross-industry coagglomeration measures for production, which is centered around 0. While there is overlap between the two distributions, the distribution for innovation and production coagglomeration measures lie clearly to the right of that for the industrial production pairs. It implies that innovation and production are more coagglomerated than the industrial production.

[6]Ellison et al. (2010) finds that the coagglomeration of the industry's production can be explained by the Marshallian forces of input sharing, labor pooling, and knowledge spillovers.

[7]Table A.1 in the Data Appendix summarizes the mean, 25th percentile, median, 75th percentile, and max coagglomeration indices of the cross-industry production for each manufacturing industry. The mean and median coagglomeration of cross-industry production are centered at zero and skewed towards positive values.

1.3 The Returns to R&D When Innovation Collocated with Production

The coagglomeration of innovation and production provides suggestive evidence on the importance of the two activities being closely located. This section provides econometric evidence on the importance of proximity to production on innovation efficiency.

1.3.1 Empirical Model

I develop and estimate an empirical model of production and innovation within the manufacturing firm to quantify the returns to R&D, primarily to assess if the returns are higher when innovation plants are located in places with more own industry production workers. The empirical model considers a manufacturing firm i with a set of active plants $j \in J_{i,t}$.[8] In each period, it determines the optimal levels of variable inputs, capital investment, R&D expenditures, and output prices for each of its plants to maximize the firm-level profits. As my focus is on exploring the spillovers from the local production to innovation, I restrict attention here to the R&D investment decisions and process of plant performance evolution and abstract from the innovation plants' decision to enter or exit. In the subsections below, I will first model the revenue functions for these production plants within the firm, then model the evolution of the plants' performance, and finally define the firms' maximization problem.

Revenue Function

Assume plant j locates at l and operates in industry k. Following Aw et al. (2011) and Bøler et al. (2015), I model its short-run marginal cost function at period t as [9]

$$\ln c_{jt} = \beta_0 + \beta_k \ln k_{jt} + \beta_w \ln w_{kt} - \psi_{jt}, \tag{1.3}$$

where k_{jt} is the capital stock of plant j at period t, w_{kt} is the wage common to all plants in industry k, and ψ_{jt} is the plant-specific productivity. The marginal costs of

[8] Plants within a manufacturing firm can be one of the following functional forms: innovation plants, production ones, and the mixed ones that both innovate and produce. For simplicity, in this paper, innovation plants refer to these that conduct R&D. It can either be a plant only conducts R&D or a mixed one. The same applies to production plants.

[9] The marginal cost function here only considers the marginal cost of production. The R&D investment decisions and the cost of innovation will be modeled in the Section 1.3.1.

production are lower for plants with higher productivities. Labor is a variable input, whereas capital is determined by the investment and capital stock in the previous period.

The product market is characterized by monopolistic competition, and the demand faced by plant j in industry k is

$$q_{jt} = Q_{kt} \left(p_{jt}/P_{kt}\right)^{-\sigma} exp\left[\zeta_{jt}\left(\sigma - 1\right)\right], \tag{1.4}$$

where $\sigma > 1$ is the constant elasticity of substitution, p_{jt} is the output price of plant j, and ζ_{jt} is a plant-specific demand shifter. The variable Q_{kt} and P_{kt} are the industry-level demand and price index.

Given the cost and demand function described above, firm i sets the optimal price p_{jt} to maximize the plant j's profits. The log revenue of plant j depends on the aggregate market conditions, the capital stock, and the plant-specific performance,

$$\ln Rev_{jt} = \gamma_0 + \ln\left(Q_{kt}P_{kt}^{\sigma}\right) + (1-\sigma)\,\beta_w \ln w_{kt} + (1-\sigma)\left(\beta_k \ln k_{jt} - z_{jt}\right) + u_{jt}, \tag{1.5}$$

where $\gamma_0 \equiv (1-\sigma)ln(\frac{\sigma}{\sigma-1}) + (1-\sigma)\beta_0$, and u_{jt} is the measurement error. Denote $z_{jt} = \psi_{jt}+\zeta_{jt}$ as the plant performance. It is an endogenous state variable that captures two sources of heterogeneity: plant-specific productivity ψ_{jt} and plant-specific demand shock ζ_{jt}. R&D investment can boost plant performance by raising its productivity or product quality. Plants with higher performance z_{jt} obtain a higher revenue.

The Impact of Innovation on Plant Performance

The performance of plant j evolves as a Markov process that depends on its own investment in R&D, the transfer of technology from other plants in the same firm and a random shock,[10]

$$z_{jt} = \alpha_0 + \alpha_1 z_{jt-1} + \alpha_2 z_{jt-1}^2 + V_{jt} + \sum_{j' \in J_t} \gamma_{jj'} V_{j't} + \eta_{jt}, \tag{1.6}$$

where V_{jt} captures the increase in plant performance through its own investment in R&D. Following Aw et al. (2011), Doraszelski and Jaumandreu (2013) and Bøler et al. (2015), I assume that the increment in plant performance at period t depends on

[10]The main focus of this paper is to explore and quantify the spillovers from the local manufacturing to innovation. Thus, I restrict my attention to the existing innovation plants and look at a subset of predetermined innovation plants. Tecu (2013) estimate an R&D location choice model to assess the importance of manufacturing on firms' innovation location choice. She finds that manufacturing plays an important role in determining innovation location, and both the internal and external linkages between innovation and production are important for the innovation plants.

its investment in R&D in the previous period. To explore and quantify the spillovers from the local production, I assume it also depends on the interaction between R&D investment $\ln r_{jt-1}$ and the employment of local production workers in its own industry $\ln(emp_{lkt-1})$. The increment in plant performance V_{jt} is written as

$$V_{jt} = \beta_1 \ln r_{jt-1} + \beta_2 \ln r_{jt-1} \ln(emp_{lkt-1}). \tag{1.7}$$

The inclusion of $\ln r_{jt-1} \ln(emp_{lkt-1})$ allows for the possibility of spillover from production to innovation within the same industry.[11] The rationale is that the local employment of its industrial production workers is a source of knowledge that can enhance plant innovation, through which it affects the plant's future performance. The empirical analysis also includes an alternative specification with discrete R&D expenditure (R&D dummy).

The term $\sum_{j' \in J_t} \gamma_{jj'} V_{j't}$ in Equation (1.7) captures intrafirm impact of R&D investment. Bilir and Morales (2016) shows that technological improvements developed in a plant can be shared with other plants within the same firm. I incorporate the sharing of proprietary technology in my empirical model by allowing the intra-firm transfer of proprietary technology across plants in the same industry.[12] $\tau_{jj'}$ captures the knowledge communication frictions when transferring technology from plant j' to plant j. Assume the decay of technology is a function of the distance between plant j and j' as follows

$$\sum_{j' \in J_t} \tau_{jj'} V_{j't} = \sum_{j' \in J_t} \left[\beta_4 + \beta_5 \ln \beta\left(dist_{jj'}\right)\right] V_{j't}. \tag{1.8}$$

The term η_{jt} in Equation (1.6) represents shocks to the performance of plant j at time t that are not anticipated at $t-1$.

Firm Optimization Problem

In each period t, firm i determines the optimal levels of labor $\mathbf{L_{it}}$, capital investment $\mathbf{I_{it}}$, R&D investment $\mathbf{R_{it}}$, and output prices $\mathbf{P_{it}}$ for each of its plants j active at time

[11] The spillover from local production to innovation considers both the internal feedbacks from its own local factories but also the manufacturing know-how learned from the other firms.

[12] In the case when the innovation-only plants conduct R&D to improve the performance of the production and the mixed plants, the spillovers are from the industry to which it provides technology. For example, if the innovation plant invests in R&D to improve the performance of the auto manufacturing plants, then the local production workers in the auto industry will be sources of expertise that can enhance the innovation plant R&D efficiency. My data sample includes firms that run businesses in more than one 3-digit NAICS industries. In this case, I restrict the transfer of technology to be only within the plants in the same 3-digit NAICS industry.

t, as well as the set of plants that will be active at $t+1$, $J_{i,t+1}$. These decisions are made based on the state vector $\mathbf{S_{it}}$ that firm i faces.

The state vector for plant j is

$$S_{ijt} = \left(z_{jt}, k_{jt-1}, w_{jt}, P_{kt}, Q_{kt}, p^k_{jt}, \ln\left(emp_{lkt-1}\right), F_{jt}\right), \tag{1.9}$$

where p^k_{jt} is the price of capital and F_{jt} is a fixed cost of operating plant j.

Decisions at period t regarding plant j's investment and employment depend on z_{jt}. Labor is a variable input, whereas capital is fixed in the short run. It's determined by the investment and capital stock in period $t-1$, according to the law of motion $k_{jt} = (1-\delta_k)k_{jt-1} + i_{jt-1}$. Decisions on plant j's investment in R&D also depend on the local employment of its own industry production workers $\ln\left(emp_{lkt-1}\right)$.

If plant j is active at period t, the profit function of plant j is

$$\Pi(s_{jt}, i_{jt}, l_{jt}, p_{jt}, r_{jt}) = \frac{1}{\sigma} Rev\left(Z_{jt}, k_{jt}, P_{j_k t}, Q_{j_k t}, w_{jt}\right) - r_{jt} - C_k\left(i_{jt}, p^k_{jt}, k_{jt}\right) - F_{jt}, \tag{1.10}$$

where $C_k(\cdot)$ is the cost function of investment in capital.

The firm i's optimization problem is

$$V(\mathbf{S_{it}}) = max_{\mathbf{X_{it}}}\left[\sum_{j\in J_{it}} \Pi\left(s_{ijt}, i_{ijt}, n_{ijt}, p_{ijt}, r_{ijt}\right) + \delta E\left(V\left(\mathbf{S_{it+1}}\right) | \mathbf{S_{it}}, \mathbf{I_{it}}, \mathbf{R_{it}}, J_{it+1}\right)\right], \tag{1.11}$$

where $\mathbf{X_{it}} = (J_{it+1}, \mathbf{L_{it}}, \mathbf{M_{it}}, \mathbf{I_{it}}, \mathbf{R_{it}}, \mathbf{P_{it}})$ is a vector of control variables, $V(\cdot)$ is the value function, $\Pi(\cdot)$ is the profit function and δ is the discount factor.

1.3.2 Estimation

To estimate the impact of local manufacturing on the returns to R&D, I proceed in two steps. First, I estimate the revenue function in Equation (1.5). Second, I estimate the Markov process governing the evolution of firm performance in Equation (1.6).

Step 1 Estimating Revenue Function Plant performance z_{jt} is likely to affect its demand for labor and capital. Therefore, OLS estimates of the revenue function suffer from the simultaneity bias. I utilize the insights from Olley and Pakes (1996) and Levinsohn and Petrin (2003) and rewrite the unobserved performance in terms of some observed variables that are correlated with it. In general, plants' usages of materials, electricity and energy depend on the level of its performance. Therefore, I rewrite the level of productivity, conditional on the capital stock, as a function of the variable input levels, i.e. $z_{jt}(k_{jt}, m_{jt}, n_{jt}, e_{jt})$. This allows me to use the expenditures on materials m_{jt}, electricity e_{jt} and energy n_{it} by the firm to control for the productivity in Equation (1.5).

The Equation (1.5) can be rewritten as

$$\ln Rev_{jt} = \gamma_0 + d_{kt} + h(k_{jt}, m_{jt}, n_{jt}, e_{jt}) + u_{jt}, \tag{1.12}$$

where the function $h(k_{jt}, m_{jt}, n_{jt}, e_{jt}) = (1-\sigma)(\beta_k \ln k_{jt} - z_{jt}(k_{jt}, m_{jt}, n_{jt}, e_{jt}))$ captures the combined effects of capital and plant performance on revenue and $d_{kt} = \ln(Q_{kt}P_{kt}^{\sigma}) + (1-\sigma)\beta_w \ln w_{kt}$ is a set of industry-time dummies capturing industry-wide demand and cost trends. I specify $h(\cdot)$ as a cubic function of its arguments and estimate Equation (1.12) by OLS.

Step 2 Estimating Performance Evolution Function In the second step, I rewrite the plant performance z_{jt} in terms of predicted $\hat{h}_{jt}$ from the first stage

$$z_{jt} = -\frac{1}{1-\sigma}\hat{h}_{jt} + \beta_k \ln k_{jt}. \tag{1.13}$$

Substituting the z_{jt} into Equation (1.6) for the performance evolution gives the estimation equation

$$\begin{aligned}
\hat{h}_{jt} \quad = \quad & -\alpha_0^* + \beta_k^* \ln k_{jt} + \alpha_1\left(\hat{h}_{jt-1} - \beta_k^* \ln k_{jt-1}\right) - \alpha_2^*\left(\hat{h}_{jt-1} - \beta_k^* \ln k_{jt-1}\right)^2 \\
& - (\beta_1^* \ln r_{j,t-1} + \beta_2^* \ln r_{j,t-1} \ln(emp_{lkt-1})) \qquad (1.14) \\
& - \sum_{j'} (\beta_3 + \beta_4 \ln(dist_{jj'}))(\beta_1^* \ln r_{j',t-1} + \beta_2^* \ln r_{j',t-1} \ln(emp_{l'kt-1})) - \eta_{jt}^*,
\end{aligned}$$

where the superscript $*$ denotes that the coefficient is multiplied by $1-\sigma$.[13]

The second stage estimation equation relates the predicted revenue to the current and lagged capital stock, the lagged predicted revenue, lagged R&D expenditure of its own, and that of other plants in the same firm, as well as the lagged local production employment of its industry.

I estimate the second stage equation with the generalized method of moments (GMM). The identification of the parameters depends on the timing assumptions. η_{jt} are the shocks to plant performance between $t-1$ and t that are unanticipated by the firm. By construction, the shocks are not correlated with the predetermined variables k_{jt-1},r_{jt-1}, $r_{j't-1}$, $\ln(emp_{jt-1})$, $\ln(emp_{j't-1})$ and $\ln(dist_{jj't-1})$. I allow the constant in the Markov process to vary by industry by including industry fixed effects (3-digit

[13]Except for α_2^*, $\alpha_2^* = \frac{\alpha_2}{1-\sigma}$.

NAICS). In total, it gives me 31 moment conditions with 28 unknown.

The α and β can be backed out given an estimate of σ. The demand elasticity σ is estimated from the total variable cost function. Given the CES demand, the total variable cost TC_{jt} can be written as a function of its total revenue Rev_{jt}:

$$TC_{jt} = (1 - \frac{1}{\sigma})Rev_{jt} + \epsilon_{jt}, \tag{1.15}$$

where the error term ϵ_{ijt} is the measurement error in total cost.

1.3.3 Data

Estimating the impact of local manufacturing on the plant innovation efficiency requires data on R&D expenditure for each innovation plant, input and output for each production plant, as well as the local employment of manufacturing industries. By combining several confidential datasets from the Census Bureau, I create a plant-level panel data on R&D investment and domestic production for US R&D firms in the manufacturing sector from 2002-2012. Details regarding data construction are documented in the Data Appendix.

Firm-level data on R&D investment are available between 1972 and 2015 from the BRDIS.[14] The BRDIS utilizes a stratified sample frame and samples firms proportionally within each stratum to their known R&D activity, or to their payroll if R&D activity is unknown. It surveys about 45,000 firms each year and includes a certainty component for firms with payroll or R&D expenditures above a certain threshold. My data sample is mainly composed of these large firms which are consistently surveyed by the BRDIS. The BRDIS asks respondents to report the total domestic R&D expenditure as well as the allocation of their spending across states.[15] The firm-level information on the geographic location of domestic R&D activities allows me to identify innovation plants and their R&D expenditure within an R&D-performing firm when supplemented by the plants location information from the LBD. In my empirical model, the innovation plants are predetermined, thus I define an establishment as an

[14]This survey was known as the Survey of Industrial Research Development (SIRD) from 1972 through 2007. The SIRD is the predecessor of the BRDIS.

[15]In the BRDIS survey, R&D is defined as planned, creative work aimed at discovering new knowledge or developing new or significantly improved goods and services. This includes 1) activities aimed at acquiring new knowledge or understanding without specific immediate commercial application or use (basic research); 2)activities aimed at solving a specific problem or meeting a specific commercial objective (applied research) 3) systematic use of research and practical experience to produce new or significantly improved goods, services, or processes (development). Thus costs for routine product testing, quality control, and technical services are not included in the R&D expenditure.

innovation plant if it has ever done R&D during the period pre-sample 1972-2001.

Plant-level production data are from the Census of Manufacturers (CMF) for years ending in 2 and 7, and from the Annual Survey of Manufactures (ASM) in other years.[16] The CMF/ASM collects production data for manufacturing establishments on their value of shipments, employment and payroll, cost of inputs, number of products, and capital. The local manufacturing employment is constructed by using data from the LBD. It's measured at the county level and on the 4-digit NAICS manufacturing industries.

In total, my data sample includes 1,500 R&D firms. Table 1.1 provides the summary statistics of the R&D firms. Within an R&D firm, the average number of innovation sites is 3.3, and the median is 2. These innovation sites locate in about 1200 counties and 400 commuting zones. The average number of production sites is 9.7, with a median of 4. On average, the innovation sites spend 12,860 thousand dollars on R&D each year.

Table 1.1 Summary Statistics of the R&D Firms

Number of R&D firms	$1500^{\dagger}$
Average number of innovation plants within a firm	3.3
Median number of innovation plants within a firm	2
Average number of production plants within a firm	9.7
Median number of production plants within a firm	4
Average R&D expenditure at innovation plants (thousands)	$12,860^{\dagger}$
Number of innovation counties	$1200^{\dagger}$
Number of innovation commuting zones	$450^{\dagger}$

Notes: This table reports the summary statistics of the R&D firms in the data sample. † indicates number is rounded to thousands or hundreds to meet the Census disclosure requirements.

1.3.4 Empirical Results

Baseline Estimates

The estimates of the empirical model described above are reported in Table 1.2. The upper panel presents the estimates of structural parameters in Equation (1.14). The

[16]The ASM is a sample survey of approximately 50,000 establishments. For sample efficiency and cost consideration, the big and important establishments in each industry are overrepresented. A number of establishments are included in the sample with certainty and the remaining establishments are sampled at a probability that is consistent with their relative importance in the industry or other key aggregations. Further details on the ASM are provided in the Data Appendix.

lower panel evaluates the mean plant performance elasticities with respect to its own R&D investment, other plants' R&D investment, and the distance between the technology receiving and providing plants, by using estimated structural parameters in the upper panel. Columns (a) and (b) use a discrete measure of R&D, while (c) and (d) use a continuous measure.

In columns (a) and (c), I first evaluate the influence of R&D investment on plant performance in a specification that does not allow the spillovers from the local manufacturing to innovation. The benchmark estimates show that a plant's performance increases in R&D investment. The results qualitatively match the recent estimates by Aw et al. (2011), Doraszelski and Jaumandreu (2013), Bøler et al. (2015) and Bilir and Morales (2016). The performance impact of a plant's own R&D investment is economically significant, and that of other plants' R&D investment is positive yet not precisely estimated.

In columns (b) and (d), I include the interaction between R&D investment and local manufacturing employment. The estimates of β_2 are positive and statistically significant, evidencing that innovating plants obtain significantly higher gains if they are located in counties with more of their own industry production workers. It suggests that there are positive spillovers from local manufacturing to innovation. All else equal, the estimates in column (b) indicate that doubling the local own industry manufacturing employment increases the impact of a plant's own investment on its performance by 21.4%. The empirical results contribute to our understanding of the role of local manufacturing in enhancing efficiency of innovation plants and it support the view put forth in Naghavi and Ottaviano (2009) that feedback from manufacturing plants is important for research labs. The complementarity between innovation and manufacturing that I find in columns (b) and (d) is also of importance from the innovation policy perspective.

The capital coefficient β_k is negative and statistically significant across all specifications. It implies that with higher capital stock, the marginal costs are lower and revenue is higher for firms. The coefficients α_1 and α_2 measure the impact of lagged productivity on current productivity. The estimates are strong and precisely estimated, indicating the impact of R&D on plant performance is persistent. The coefficients β_3 and β_4 quantify knowledge communication frictions when transferring technology within the firm across plants, and the lower panel calculates the mean performance elasticity with respect to distance. The estimates show that the longer the distance between the technology receiving and providing plants, the larger the performance losses will be. I check the validity of the instruments with an overidentification test and the p-values from the test are listed below the estimates in each column.

To retrieve the structural parameter α and β, I estimate σ from Equation (1.15). The estimates of $1 - \frac{1}{\sigma}$ is 0.6618 with the standard error of 0.0111.

Table 1.2 GMM Estimates: Baseline

		Continuous R&D		Discrete R&D	
		(a)	(b)	(c)	(d)
β_k	Log(capital)	-0.4073***	-0.4087***	-0.4073***	-0.4083***
		(0.0043)	(0.0043)	(0.0043)	(0.0042)
α_1	Performance$_{t-1}$	0.9088***	0.9120***	0.9093***	0.9114***
		(0.0112)	(0.0115)	(0.0112)	(0.0114)
α_2	Performance$^2_{t-1}$	0.0969***	0.0979***	0.0972***	0.0979***
		(0.0084)	(0.0085)	(0.0083)	(0.0084)
β_1	Log(R&D$_{t-1}$)	0.0014***	-0.0005	0.0118***	-0.0086*
		(0.0002)	(0.0005)	(0.0021)	(0.0045)
β_2	Log(R&D$_{t-1}$)$\times$ Log(Emp$_{t-1}$)		0.0003***		0.0031***
			(0.0001)		(0.0008)
β_3	Constant	0.3216	0.2804	0.379	0.2187
		(0.2567)	(0.2464)	(0.2628)	(0.2467)
β_4	Log(Dist.)	-0.0763*	-0.0783**	-0.0788*	-0.0631*
		(0.0398)	(0.0375)	(0.0407)	(0.0375)
Own Plant R&D Elasticity		0.0014***	0.0014***	0.0118***	0.0124***
		(0.0002)	(0.0002)	(0.0021)	(0.0019)
Other Plant R&D Elasticity		0.0001	0.0000	0.0016	-0.0000
		(0.0002)	(0.0001)	(0.0016)	(0.0011)
Distance Elasticity		-0.0002*	-0.0002**	-0.0014*	-0.0009*
		(0.0001)	(0.0001)	(0.0007)	(0.0006)
Overidentification Test (p value)		0.81	0.31	0.82	0.34
Observations		106,000	106,000	106,000	106,000
Industry Effects		Yes	Yes	Yes	Yes

Notes: The upper panel reports the GMM estimates of structural parameters in Equation (1.14). Columns (a) and (c) presents the estimates of the benchmark specification. Robust standard errors in parentheses are clustered by county. Each specification reports the p-value for the overidentification restrictions test (Hansen, 1982). The lower panel evaluates the mean plant performance elasticities by using estimated structural parameters in the upper panel. R&D is measured as log(1+R&D) in columns (a) and (b), and a binary variable in columns (c) and (d). *** denotes 1% significance, ** 5%, and * 10%.

Instrumenting Local Manufacturing Employment with Predicted Employment

A potential concern is the local employment might be positively correlated with the unobserved shocks to plant performance. To address this concern, I instrument the local manufacturing employment with a predicted one. The predicted local employment is constructed by interacting the initial local employment shares with national growth of industry employment *a la* Bartik (1991).

Using the pre-sample year 1997 as the base year, I predict the local manufacturing employment for each 4-digit NAICS manufacturing industry by interacting the initial shares of 6-digits NAICS subindustry in year 1997 with the national growth of the 6-digits NAICS subindustry employment. The formula is as follows

$$pemp_{lk_{4d}t} = \sum_{k_{6d} \in k_{4d}} s_{lk_{6d},1997} \times g_{k_{6d},t}, \tag{1.16}$$

where $pemp_{lk_{4d}t}$ is the predicted manufacturing employment at county l in 4-digit NAICS industry k at time t, $s_{lk_{6d},1997}$ is the initial share of each 6-digit NAICS subindustry within the 4-digit NAICS industry k at county l in the year 1997 and the $g_{k_{6d},t}$ is the national growth rate of each 6-digit NAICS subindustry between the year 1997 and time t.

Table 1.3 reports the estimates when instrumenting local manufacturing employment with a predicted one. It gives similar estimates on the structural parameters as our baseline estimation in Table 1.2. The impact of local manufacturing on innovation efficiency β_2 is slightly lower.

Robustness Checks

Spillovers from the Local R&D A potential concern is that other than the spillovers from the local manufacturing to the plant innovation, there are also spillovers from the local R&D activities of the other firms. As a robustness check, I estimate a specification that controls the spillovers from the local innovation. The increment in plant performance in Equation (1.7) is written as

$$V_{jt} = \beta_1 \ln r_{jt-1} + \beta_2 \ln r_{jt-1} \ln\left(emp_{lkt-1}\right) + \beta_5 \ln r_{jt-1} \ln(rdemp_{lt-1}). \tag{1.17}$$

where $\ln(rdemp_{lt-1})$ is the local R&D employment. Due to the data availability, I measure the local R&D employment at the commuting zone level.[17] Local R&D employment is measured as the total employment in the NAICS 5417 *Scientific Research and Development Services* industry.

[17]The plant-level data on R&D employment from the BRDIS is highly correlated with the R&D expenditure. Since the R&D investment is lumpy, local R&D employment measured by data from the BRDIS is lumpy as well. Instead, I measure the local R&D employment by using data from the LBD and measured it as the employment in the *Scientific Research and Development Services* industry. The county-level R&D employment is limited, and thus I measure the local R&D employment at the county level.

Table 1.3 GMM Estimates: Instrumenting Local Employment with Predicted Emp.

		Continuous R&D		Discrete R&D	
		(a)	(b)	(c)	(d)
β_k	Log(capital)	-0.4073***	-0.4088***	-0.4073***	-0.4083***
		(0.0043)	(0.0042)	(0.0043)	(0.0042)
α_1	Performance$_{t-1}$	0.9088***	0.9125***	0.9093***	0.9118***
		(0.0112)	(0.0115)	(0.0112)	(0.0113)
α_2	Performance$^2_{t-1}$	0.0969***	0.0986***	0.0972***	0.0985***
		(0.0084)	(0.0084)	(0.0083)	(0.0084)
β_1	Log(R&D$_{t-1}$)	0.0014***	-0.0001	0.0118***	-0.0051
		(0.0002)	(0.0005)	(0.0021)	(0.0044)
β_2	Log(R&D$_{t-1}$)$\times$ Log(Emp$_{t-1}$)		0.0002**		0.0025***
			(-0.0001)		(0.0008)
β_3	Constant	0.3216	0.3638	0.379	0.3016
		(0.2567)	(0.2546)	(0.2628)	(0.2600)
β_4	Log(Dist.)	-0.0763*	-0.0905**	-0.0788*	-0.0758*
		(0.0398)	(0.0392)	(0.0407)	(0.0398)
Own Plant R&D Elasticity		0.0014***	0.0013***	0.0118***	0.0118***
		(0.0002)	(0.0002)	(0.0021)	(0.0019)
Other Plant R&D Elasticity		0.0001	0.0001	0.0016	0.0004
		(0.0002)	(0.0001)	(0.0016)	(0.0012)
Distance Elasticity		-0.0002*	-0.0002**	-0.0014*	-0.0011*
		(0.0001)	(0.0001)	(0.0007)	(0.0006)
Overidentification Test (p value)		0.81	0.25	0.82	0.34
Observations		106,000	106,000	106,000	106,000
Industry Effects		Yes	Yes	Yes	Yes

Notes: The upper panel reports the GMM estimates of structural parameters in Equation (1.14), instrumenting local employment with the predicted one. Columns (a) and (c) presents the estimates of the benchmark specification. Robust standard errors in parentheses are clustered by county. Each specification reports the p-value for the overidentification restrictions test (Hansen, 1982). The lower panel evaluates the mean plant performance elasticities by using estimated structural parameters in the upper panel. R&D is measured as log(1+R&D) in columns (a) and (b), and a binary variable in columns (c) and (d). *** denotes 1% significance, ** 5%, and * 10%.

Table 1.4 reports the estimates for the robustness check. The inclusion of the spillovers from the R&D activities to the plants' innovation has a negligible impact on my results. The estimated coefficient β_5 for the spillovers from local R&D is not significantly different from zero. Thus, there is no evidence suggesting the spillovers from local R&D to plant innovation.

The Direct Impact of Local Employment on Plant Performance Another concern is that the local employment might have a direct impact on plant performance. To account for this possibility, I consider a specification that includes the employment levels in the performance increment function as follows,

$$V_{jt} = \beta_1 \ln r_{jt-1} + \beta_2 \ln r_{jt-1} \ln \left(emp_{lkt-1}\right) + \beta_6 \ln \left(emp_{lkt-1}\right). \tag{1.18}$$

Accounting for the direct effect of local manufacturing employment does not change the conclusions from my baseline estimation. The local manufacturing employment has no direct impact on plant performance, and it only affects the performance through interacting with the R&D investment.

Table 1.4 GMM Estimates: Robustness

		Continuous R&D		Discrete R&D	
		(a)	(b)	(c)	(d)
β_k	Log(capital)	-0.4085***	-0.4087***	-0.4081***	-0.4083***
		-0.0043	-0.0042	-0.0042	-0.0042
α_1	Performance$_{t-1}$	0.9115***	0.9123***	0.9110***	0.9117***
		-0.0114	-0.0114	-0.0113	-0.0113
α_2	Performance$^2_{t-1}$	0.0977***	0.0985***	0.0977***	0.0985***
		-0.0084	-0.0084	-0.0084	-0.0084
β_1	Log(R&D$_{t-1}$)	-0.0002	-0.0002	-0.0055	-0.006
		-0.0006	-0.0006	-0.0052	-0.0051
β_2	Log(R&D$_{t-1}$)$\times$ Log(Emp$_{t-1}$)	0.0003***	0.0002**	0.0034***	0.0022***
		-0.0001	-0.0001	-0.0008	-0.0003
β_5	Log(R&D$_{t-1}$)$\times$ Log(RDemp$_{t-1}$)	-0.0001		-0.0008	
		-0.0001		-0.0007	
β_6	Log(Emp$_{t-1}$)		0.0001		0.0004
			-0.0002		-0.0003
β_3	Constant	0.2889	0.3574	0.2296	0.2791
		-0.2339	-0.2436	-0.2349	-0.2303
β_4	Log(Dist.)	-0.0793**	-0.0872**	-0.0644*	-0.0684*
		-0.036	-0.038	-0.036	-0.0357
Overidentification Test (p value)		0.32	0.25	0.35	0.34
Observations		106,000	106,000	106,000	106,000
Industry Effects		Yes	Yes	Yes	Yes

Notes: This table reports GMM estimates corresponding to variants of Equation (1.14). Columns (a) and (c) allow for spillovers from the local R&D activities to innovation. Columns (b) and (d) incorporate the direct impact of local manufacturing employment on plant performance. Robust standard errors in parentheses are clustered by county. Each specification reports the p-value for the overidentification restrictions test (Hansen, 1982). R&D is measured as log(1+R&D) in columns (a) and (b), and a binary variable in columns (c) and (d). *** denotes 1% significance, ** 5%, and * 10%.

1.4 A Model of Innovation and Production

Motivated by the facts presented in section 1.2 and the empirical evidence on the positive spillovers from local manufacturing to innovation in section 1.3, I extend the trade and multi-region production model developed by Arkolakis et al. (2018) with two key modifications: (i) allowing firms to increase productivity through R&D; and (ii) incorporating the spillovers from local manufacturing to innovation.

1.4.1 Setup

Consider an economy with $n = 1 \cdots N$ regions. There are $\bar{L}_n$ workers in region n. Workers with CES preferences consume a continuum of goods indexed by $\omega \in \Omega$:

$$U_n = \left(\int_{\omega \in \Omega} q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega \right)^{\frac{\sigma}{\sigma-1}}. \tag{1.19}$$

There are two activities in the economy: innovation and production. Workers possess heterogeneous abilities in these activities. Some are good at production while others are good at research. Their abilities in these two activities are characterized by the efficiency units of labor with which they are endowed, $\boldsymbol{E} = (E^e, E^p)$. E^e denotes the endowment of efficiency units of labor which can be supplied to innovation activities, and E^p denotes the endowment that can be supplied to production. Assume that $E^e = u_e / \Gamma(1 - 1/\kappa)$ and $E^p = u_p / \Gamma(1 - 1/\kappa)$, with u_e and u_p drawn independently from the distribution $F(u) = \exp[-u^{-\kappa}]$, where $\kappa > 1$ and $\Gamma(\cdot)$ is the Gamma function.

Workers are immobile across different regions but mobile across innovation and production activities within each region.[18] Wage per efficiency unit of innovation labor is w_n^e, and per efficiency unit of production labor is w_n^p. Workers will choose to work in innovation activities if $E^e w_n^e \geq E^p w_n^p$, otherwise they will choose to work in production. Given the wages w_n^e and w_n^p, the supply of labor units to innovation and production activities in region i are given by

$$L_n^e = \bar{L}_n \left[1 + \left(\frac{w_n^e}{w_n^p} \right)^{-\kappa} \right]^{1/\kappa - 1}, \tag{1.20}$$

[18] There is limited empirical evidence of geographic mobility. Caliendo et al. (2015) find that only 2% of the US population moves across states in a year. Autor et al. (2013) find that trade shocks induced only small population shifts across regions in the US. In the appendix, I consider an extension of the benchmark model where workers can move across regions.

and

$$L_n^p = \bar{L}_n \left[1 + \left(\frac{w_n^p}{w_n^e} \right)^{-\kappa} \right]^{1/\kappa - 1}. \tag{1.21}$$

Labor units supplied to innovation and production activities depend on the relative wage $\frac{w_n^e}{w_n^p}$. With a finite κ, workers are heterogeneous in their productivity across activities. The change of relative wage will lead to the expansion of one activity and the contraction of the other. With $\kappa \to \infty$, workers are homogeneous and perfectly mobile across activities. There is no mobility across activities when $\kappa \to 1$.

1.4.2 The R&D Firm's Problem

A firm born in region i only conducts R&D in its birth region.[19] New technologies created through R&D can be used in multi-region production (MP). MP occurs when the technology from region i is used for production in region l. The firm's productivity in region l is denoted as z_l. Engaging in MP incurs a productivity loss γ_{il} due to the transfer of technology. γ_{il} is an iceberg cost, with $\gamma_{il} > 1$ and $\gamma_{ll} = 1$. There is a fixed cost F_n in units of labor and an iceberg trade cost τ_{ln} when selling a variety produced in l to n. Labor is the only input of production, and therefore the marginal cost of a variety from i produced in region l to serve market n is $c_{iln} = \frac{\gamma_{il} w_l^p \tau_{ln}}{z_l}$. Given the CES preference, firms will set the prices $p_{iln} = \tilde{\sigma} c_{iln}$, where $\tilde{\sigma} = \frac{\sigma}{\sigma - 1}$ is the markup over the marginal cost. To enter market n, the variable profits earned in market n should be able to cover the fixed cost $w_n F_n$, and thus the unit cost of production needs to be lower than $c_n^* = \left(\frac{\sigma w_n^p F_n}{X_n} \right)^{1/(1-\sigma)} \frac{P_n}{\tilde{\sigma}}$, where X_n is the total expenditure in region n and $P_n = \left(\int_{\omega \in \Omega} p_n^{1-\sigma} d\omega \right)^{1/(1-\sigma)}$ is the aggregate price index in region n.

Innovation

An R&D firm at region l can invest $\bar{f}_i$ efficiency units of innovation labor to come up with a new variety (product innovation).[20] The new variety can be produced in regions different from where it is created. Assume that the vector of productivity at each potential production site $\mathbf{z} = (z_1, z_2, \ldots, z_N)$ is randomly drawn from the multivariate

[19] For the origin-production-market triplet below, I use index i to denote the source of idea, index l to denote the location of production and index n to denote the product market.

[20] To line up with my empirical model, innovation site within an R&D firm is predetermined.

Pareto distribution

$$G_i\left(z_1,\ldots,z_N\right)=1-\left(\sum_{l=1}^{N}\left[\left(\bar{T}_i^e v\right)T_l^p z_l^{-\theta}\right]^{\frac{1}{1-\rho}}\right)^{1-\rho} \tag{1.22}$$

with support $z_l \geq \sum_l \left[T_{il}^{1/(1-\rho)}\right]^{\frac{1-\rho}{\theta}}$, $\theta > max\left(1,\sigma-1\right)$ and $\rho \in [0,1)$.[21] The shape parameter θ controls the heterogeneity across realizations of different productivity vectors, and the correlation parameter ρ controls the correlation of the elements within the productivity vector.

The scale parameter $\left(\bar{T}_i^e v\right)T_l^p$ determines the average productivity of varieties created in i and produced in l. T_l^p is the productivity in production at region l. $\left(\bar{T}_i^e v\right)$ determines the quality of ideas created in region i, and can be thought of as the productivity in innovation. $\bar{T}_i^e$ is the fundamental productivity in innovation at location i. The R&D firm can also invest in R&D to increase productivity, and thus lower the marginal cost of production (process innovation). v is the level of process innovation chosen by the firm.

To achieve v level of process innovation, the R&D firm incurs a cost of v^{β_i} $(\beta_i > 1)$ efficiency units of innovation labor. The region-specific parameter β_i reflects the capability of a region in fostering innovation. A lower β_i indicates that region i is better at fostering innovation. In this case, the marginal cost of process innovation will be lower and the returns to R&D are higher. In section 1.3, I find that returns to R&D are higher when innovation plants are located in a region with more manufacturing employment. In other words, regions with a higher employment in production are better at fostering innovation. Thus, I impose the following assumption on the region-specific parameter β_i.

Assumption: $\beta_i = \frac{1}{g\left(L_i^p\right)}$ is an decreasing function in L_i^p.

Regions with a higher level of employment in production are better at fostering innovation. In this case, the cost of process innovation $v^{\frac{1}{g\left(L_i^p\right)}}$ is lower. Denote $T_i^e = \bar{T}_i^e v$ as the post process innovation productivity. With a higher T_i^e, the new varieties created in region i will have higher productivity at all potential production sites.

Production

The R&D firm from region i faces a tradeoff between market proximity and production capability when choosing where to produce for market n. It can either locate in a

[21] See Arkolakis et al. (2017) for detail information of the distribution properties and boundary conditions.

region closer to the market (a smaller τ_{ln}) or in a region with lower production cost (a lower $\frac{w_l^p \gamma_{il}}{z_l}$). Given the multivariate Pareto distribution of the productivity vector, the probability that a variety created from i serving market n through region l is

$$\psi_{iln} \equiv Pr\left(argmin_l C_{iln} = l | min_l C_{iln} \leq c_n^*\right) = \frac{\left(T_l^p \left(\gamma_{il} w_l^p \tau_{ln}\right)^{-\theta}\right)^{1/(1-\rho)}}{\sum_m \left(T_m^p \left(\gamma_{im} w_m^p \tau_{mn}\right)^{-\theta}\right)^{1/(1-\rho)}}. \tag{1.23}$$

The probability depends on the production capacities of region l relative to other regions. The numerator represents a region's production capacity, which depends on the region's productivity in production T_l^p, the proximity to technology γ_{il} and market τ_{ln}, and the wage of production labor w_l^p. The denominator in equation is the sum of all potential production sites' capacities, and can be thought of as the access to production for firms from region i.

Given the access to production, the probability of R&D firm from i serving market n at a cost lower than c, for $c \leq c_n^*$, is

$$Pr\left(min_l C_{iln} \leq c\right) = T_i^e \left[\left(\sum_m T_m^p \left(\gamma_{im} w_m^p \tau_{mn}\right)^{-\theta}\right)^{1/(1-\rho)}\right]^{1-\rho} \theta c^{\theta-1}. \tag{1.24}$$

Denote $\Phi_{in}(v) = \left(\bar{T}_i^e v\right) \left[\left(\sum_m T_m^p \left(\gamma_{im} w_m^p \tau_{mn}\right)^{-\theta}\right)^{1/(1-\rho)}\right]^{1-\rho}$ as the market potential of firms from region i in serving market n. With higher productivity in innovation and greater access to production, firms from i gain a higher market potential in n.

Optimal Level of Process Innovation

The R&D firms choose the optimal level of process innovation to maximize their profits. Given the probability of serving market n by firms from i in Equation 1.24 and the probability of producing in region l in Equation (1.23), the expected sales of a firm from i serving market n through region l can be written as

$$E(x_{iln}) = \phi_{iln} \Phi_{in}(v) X_n \tilde{\sigma}^{1-\sigma} P_n^{\sigma-1} \int_0^{c_n^*} \theta c^{\theta-\sigma} dc, \tag{1.25}$$

and the total expected profits net of innovation costs for a firm innovating in i and conducting v level of process innovation is

$$E\left[\pi_i(v)\right] = \frac{\sigma-1}{\theta-\sigma+1} \sum_n \Phi_{in}(v) \left(\frac{X_n P_n^{\sigma-1}}{w_n^p F_n} \frac{\sigma}{\tilde{\sigma}^{1-\sigma}}\right)^{\frac{\theta}{\sigma-1}} w_n^p F_n - w_i^e \bar{f}_i - w_i^e v^\beta. \tag{1.26}$$

The first-order condition for the choice of the optimal level of process innovation and the zero expected profit condition due to the free entry yields $v = \left(\frac{1}{1-g(L_i^p)}\bar{f}_i\right)^{g(L_i^p)}$. The optimal level of process innovation is an increasing function in its production labor L_i^p. R&D firms locating in regions with higher employment in production will choose a higher level of process innovation.

1.4.3 Aggregation

The R&D firm from region i will spend f_i efficiency units of labor on innovation, which is the sum of its expenditures on product innovation $\bar{f}_i$ and process innovation $\frac{1}{g(L_i^p)}\bar{f}_i$. In region i, L_i^e efficiency units of labor are allocated to innovation and the measure of varieties created in region i is $M_i = \frac{L_i^e}{f_i}$. The total sales of varieties created in i serving the market n through region l can be written as

$$X_{iln} = M_i x_{iln} = \phi_{iln} \frac{M_i \Phi_{in}}{\sum_k M_k \Phi_{kn}} X_n. \tag{1.27}$$

Given the aggregate trilateral technology and trade flows X_{iln}, the total value of varieties produced in region l is denoted as $Y_l \equiv \sum_{i,n} X_{iln}$ and the total expenditure in region n is $X_n \equiv \sum_{i,l} X_{iln}$. X_{iln} can be used to construct three sets of aggregate bilateral shares:
the expenditure shares

$$\lambda_{in}^E \equiv \frac{\sum_l X_{iln}}{X_n} = \frac{M_i \Phi_{in}}{\sum_k M_k \Phi_{kn}}, \tag{1.28}$$

the trade shares

$$\lambda_{ln}^T \equiv \frac{\sum_i X_{iln}}{X_n} = \sum_i \phi_{iln} \lambda_{in}^E, \tag{1.29}$$

and the MP shares

$$\lambda_{il}^M \equiv \frac{\sum_n X_{iln}}{Y_l} = \frac{\sum_n \phi_{iln} \lambda_{in}^E X_n}{Y_l}. \tag{1.30}$$

1.4.4 Equilibrium

Given the measure of varieties created in region i and the profits earned by the R&D firm in Equation (1.26), the total profits earned by firms from region i can be written as $\Pi_i = M_i \pi_i = \eta \sum_n \lambda_{in}^E X_n - w_i^e L_i^e$, where $\eta = \frac{\sigma-1}{\theta\sigma}$. Zero profit condition implies that labor market clearing condition for innovation workers can be written as follows,

$$\eta \sum_n \lambda_{in}^E X_n = w_i^e L_i^e. \tag{1.31}$$

The labor demand for production in region l equals the total output Y_l minus the profits associated with the output $\frac{Y_l}{\sigma}$, which gives $\frac{\sum_n \lambda_{ln}^T X_n}{\tilde{\sigma}}$. The labor demand for serving the market (entry) is $(1 - \eta - \frac{1}{\tilde{\sigma}})X_l$, which depends on the total consumption in region l. The labor marketing clearing condition for innovation workers can be written as follows,

$$\frac{\sum_n \lambda_{ln}^T X_n}{\tilde{\sigma}} + (1 - \eta - \frac{1}{\tilde{\sigma}})X_l = w_l^p L_l^p. \tag{1.32}$$

Following Dekle et al. (2007), this model allows for aggregate trade and MP imbalances via exogenous cross region transfer Δ_i with $\sum_i \Delta_i = 0$. The budget balance condition can then be written as:

$$w_i^p L_i^p + w_i^e L_i^e + \Delta_i = X_i. \tag{1.33}$$

Equations (1.31) and (1.32) can be written in terms of wages by substituting L_i^e and L_i^p using Equations (1.20) and (1.21), and X_i using Equation (1.33). Equilibrium wages can be obtained by solving a system of $2N$ equations.

1.4.5 Calibration

I take the model to the data in year 2012, and restrict my analysis to 48 states in the US and the rest of the world (ROW) for which I have good data for trade, output and multi-region production.[22] The parameters to be calibrated in the model are the multi-variate Pareto distribution parametrs ρ and θ; the Frechet distribution parameter κ; the elasticity of substitution σ; the entry cost of innovation f_i; the parameters that determine bilateral trade and MP cost τ_{ln} and γ_{il}; the productivity parameters T_i^e and T_l^p; and regions' capability in fostering innovation $g(L_i^p)$.

To calibrate these parameters, I construct data on the bilateral trade and MP flows, the number of varieties in each region and the endowment of equipped labor in each region. For domestic trade, I aggregate the firm-level manufacturing trade flow data from the Commodity Flow Survey (CFS) to the state level to get the bilateral trade flows between the 48 states. I get the trade flows between each state and the ROW from Longitudinal Firm Trade Transaction Database (LFTTD). For the trade flows

[22] Alaska, Hawaii and D.C. are excluded from my data sample.

within the ROW $X_{row,row}$, I take it from the World Input-Output Database. With these data, I construct the 49×49 matrix of trade shares λ^T_{ln}, a vector of aggregate expenditure on manufacturing goods $X_n = \sum_l X_{ln}$, and a vector of aggregate output in manufacturing $Y_l = \sum_n X_{ln}$.

The empirical counterpart of bilateral MP flows from region i to region l is defined as the output in region l that is using the technology from region i. To construct the domestic MP flows, I link the BRDIS with Census of Manufactures (CMF). The data from the BRDIS allow me to identify the origin state of technology within a firm and the data from the CMF provide me with the information of output. I measure the outward MP flows from the US to the ROW as the output of the R&D firms' foreign affiliates. To measure the inward MP flows from the ROW to each state, I link the Survey of Business Owners (SBO) with CMF to identify the foreign firms' operations in each state. I take the MP flows from ROW to ROW from Arkolakis et al. (2018). In this way, I obtain the 49×49 matrix of trade shares λ^M_{il}.

The number of varieties created in each state and the ROW M_i are measured by linking the BRDIS (or SBO) with CMF. The data from the BRDIS and SBO allow me to identify the origin of technology for firms' production, and the data from the CMF provide information on the number of products.

The aggregate labor endowments for the US and the ROW are measured by using total equipped labor data from Penn World Table (PWT), multiplied by the share of employment in manufacturing sector from the UNIDO. The equipped labor in each state is then constructed by multiplying the share of employment in each state by the total equipped labor measured from the PWT.

Calibrated Parameters and Targeted Moments

Table 1.5 summarizes the calibrated parameters and the targeted moments in the data. I take the multi-variable Pareto shape parameter $\theta = 4.5$, correlation parameter $\rho = 5.5$, and the elasticity of substitution from $\sigma = 4$ from Arkolakis et al. (2018). I set Frechet distribution parameter $\kappa = 3$ following Hsieh et al. (2013) and Lagakos and Waugh (2013).

The rest of the parameters are calibrated following Arkolakis et al. (2018). First, I implement an extended version of Head and Ries (2001) approach to estimate the bilateral trade and MP costs. Given the data on trade, MP shares and the total consumption X_n, the model determines all the trilateral trade flows X_{iln}.[23] Assume

[23]See appendix for the proofs.

that both the trade costs and MP costs are symmetric, I compute the matrices $\hat{\tau}_{ln} = \left(\sqrt{\frac{X_{inn}X_{ill}}{X_{iln}X_{inl}}}\right)^{\frac{1-\rho}{\theta}}$ and $\hat{\gamma}_{il} = \left(\sqrt{\frac{X_{iin}X_{lln}}{X_{iln}X_{lin}}}\right)^{\frac{1-\rho}{\theta}}$.

Given the estimated matrices of trade and MP cost, I set parameters T_i^e, T_i^p and f_i to match MP deficit $\sum_l \lambda_{il}^M Y_l - Y_i$, the total output Y_i and the number of varieties respectively.

Table 1.5 Calibrated Parameters and Data Targets

Parameters	Value	Description	Target/Source
σ	4	elasticity of substitution	Arkolakis et al. (2018)
κ	3	Frechet shape parameter	Lagakos and Waugh (2013), Hsieh et al. (2013)
θ	4.5	MVP shape parameter	Arkolakis et al. (2018)
ρ	0.55	MVP correlation parameter	Arkolakis et al. (2018)
$\bar{f}$		innovation entry cost	number of varities
τ_{ln}		trade cost	bilateral trade shares
γ_{il}		MP cost	bilateral MP shares
T_i^P		productivity in production	gross output
T_i^e		productivity in innovation	MP deficit

Notes: This table summarizes the calibrated parameters and the targeted moments in the data.

Mapping Revenue Equation to the Empirical Revenue Function

The region's capability in fostering innovation $g(L_i^p)$ is estimated by mapping the revenue function from the theoretical model to the empirical model. Given the expected sales of firm i serving the market n through l in Equation (1.25), the log expected revenue of firms from i with innovation level $v = \left(f_i - \bar{f}_i\right)^{g(L_i^p)}$ choosing region l to serve all potential markets can be written as[24]

$$\ln(x_{il}) - \frac{1}{1-\rho}\ln\left(T_l^p w_l^{-\theta}\right) - \ln(D_{il}) - \ln T_i^e + \frac{\theta}{1-\rho}\ln(\gamma_{il}) = g\left(L_i^P\right)\ln\left(f_i - \bar{f}_i\right), \tag{1.34}$$

where $D_{il} = \sum_n \frac{\sum_l T_l^p \left(\gamma_{il} w_l^p \tau_{ln}\right)^{-\theta}}{\sum_k M_k \Phi_{kn}} (X_n \tau_{ln})$ can be thought of as the aggregate demand for the production plants using technologies from region i to serve the markets, and $\tilde{f}_i - f_i = \frac{1}{1-g(L_i^p)} f_i$ is the R&D expenditure that firm i spends on process innovation.

The log revenue in the empirical model is a function of the aggregate market conditions, the capital stock and the plant-specific performance, as in Equation (1.5).

[24]Firms from region i will choose the optimal level of process innovation $v = \left(\frac{1}{1-g(L_i^p)}\bar{f}_i\right)^{g(L_i^p)} = \left(f_i - \bar{f}_i\right)^{g(L_i^p)}$.

Plant-specific performance evolves over time, and depends on its own investment in R&D, as well as the transfer of technology from other R&D plants in the same firm. The performance evolution function is general enough to cover different organization structures of production and innovation within an R&D firm. For example, it allows more than one innovation plant within an R&D firm to provide technology for the production plants. The theoretical model is a simplification of what we can observe in the real data, assuming there is only one innovation plant within an R&D firm. To map the log revenue function to the empirical model, I simplify the performance evolution function by considering the case that there is only one innovation plant within an R&D firm, and I look at the long-run impact of innovation on plant performance.[25]

The log revenue function from the empirical model can then be written as

$$\tau_{jj'}^{-1}\left(\ln Rev_j - d_{j_k} - \frac{\sigma-1}{1-\alpha_1}\alpha_0\right) = \frac{\sigma-1}{1-\alpha_1}\left(\beta_1 \ln r_{j'} + \beta_2 \ln r_{j'} \ln\left(emp_{j'_{l'k}}\right)\right), \quad (1.35)$$

where plant j is the technology-receiving plant and plant j' is the innovation plant.[26]

Comparing the revenue function from the theoretical model in Equation (1.34) with that from the empirical model in Equation (1.35), the left hand side of both equations gives the production plants' revenue after controlling for region-specific characteristics in production, the aggregate demand, the fundamental productivity and the loss of technology from the transferring. The region-specific productivity in production is captured by the term $\ln\left(T_l^p w_l^{-\theta}\right)$ in the theoretical model, and the empirical counterpart is d_{j_k}. The aggregate demand depends on D_{il}, which is captured by d_{j_k} in equation 1.35. Both equations relate revenue to the long-run predicted productivity. $\bar{T}_i^e$ determines the fundamental productivity in innovation in the model and $\frac{\sigma-1}{1-\alpha_1}\alpha_0$ captures the long-run effects of innovation on productivity. The friction of technology transfer is captured by γ_{il} theoretically and by $\tau_{jj'}$ empirically.

The right hand sides of both equations relate the production plant's revenue to the investment in R&D. In the theoretical model, the plant's revenue depends on the investment in R&D at the innovation plant, as well as the regions' capacity in fostering innovation. The capacity is an increasing function of the local manufacturing employment, denoted as $g(L_i^p)$. Empirically, the plant's revenue depends on the investment in R&D $\frac{\sigma-1}{1-\alpha_1}\left(\beta_1 \ln r_j\right)$ and the spillovers from the local manufacturing $\frac{\sigma-1}{1-\alpha_1}\left(\beta_2 \ln r_{j'} \ln\left(emp_{j'_{l'k}}\right)\right)$. Given that the estimate of β_1 in Equation (1.7) is not

[25]The long run effect of R&D investment on plant performance is captured by α_1 and α_2. More than 90% of current period can be explained by its performance in the previous period. For simplification, I only consider the first-order effects of R&D investment on plant long-run performance.

[26]I surpress the constant term $\gamma_0 = log(\frac{\theta}{\theta-\sigma+1})$ in the equation.

significantly different from zero, I will mainly focus on the spillover effects from the local manufacturing. Assume innovation capacity $g(L_i^p) = \mu \ln L_i^p$, where μ captures the strength of the spillovers from the local manufacturing to innovation. The spillover strength parameter takes the value $\mu = \frac{\hat{\sigma}-1}{1-\hat{\alpha}_1}\hat{\beta}_2$, where the estimates are from Table 1.3 column (b). With $\hat{\alpha}_1 = 0.9125$, $\hat{\beta}_2 = 0.0002$ and $\hat{\sigma} = 2.9568$, the strength parameter is estimated to be $\mu = 0.005$.

The benchmark estimation in the empirical section also considers the case of no spillovers from the local manufacturing to innovation. In this case, the plant's revenue only depends on the innovation plant's investment in R&D. To shut down the spillovers from the local manufacturing in the model, I assume the local innovation capacity is a constant and it equals $g(\cdot) = \frac{\hat{\sigma}-1}{1-\hat{\alpha}_1}\hat{\beta}_1$, where the estimates are taken from Table 1.3 column (a). With $\hat{\alpha}_1 = 0.9088$, $\hat{\beta}_2 = 0.0014$ and $\hat{\sigma} = 2.9568$, the constant local innovation capacity is estimated to be $g(\cdot) = 0.03$. In the counterfactural exercises, I will compare the changes of innovation efficiency in cases where there are spillovers to these cases where there are no spillovers.

1.4.6 Counterfactural Exercises

The increase in US imports from China has asymmetric impacts across regions. Autor et al. (2013) shows that labor markets with greater exposure to the increase in import competition from China experienced a larger decrease in manufacturing employment. In this section, I will use the China import competition to quantify the effects of production reallocation on the local innovation efficiency.

Identifying the Trade Shocks

Given that not all the observed changes in US imports from China are the results of a change in Chinese productivity, I replicate the procedure in Autor et al. (2013) to identify the supply-driven components of Chinese imports. I compute the predicted changes in the US imports from China *a la* Bartik. The predicted changes are the inner product of the initial US imports in each sector and the sectoral growth of Chinese imports in eight other developed countries,[27]

$$\Delta IMP_{China \to US} = \sum_k IMP_{1997}^k \times g_{China \to OTH}^k, \tag{1.36}$$

[27] The eight other developed countries are Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, and Switzerland.

where IMP_{1997}^{k} is the US imports from China in 1997, $g_{China\to OTH}^{k}$ is the growth of Chinese imports in eight other developed countries in sector k, and $\Delta IMP_{China\to US}$ is the changes in imports during the period 1997-2012.

I model the rise of China as the productivity shocks to the ROW ΔT_{ROW}^{p}, and use the predicted changes in the US imports from China to quantify the size of the productivity shocks. I calibrate these shocks such that the simulated changes in aggregate expenditure shares on goods from the ROW match the change in these expenditure shares that is driven by the rise of China during 1997-2012.

The Impact of Production Reallocation on Innovation

In this section, I will remove the China shock to see how that will affect the innovation and production across all states.

Figure 1.3 plots the percentage changes in innovation and production labor for each state when removing the import competition from China. In this case, each state will have a larger employment in production but less employment in innovation. That is to say, most states will be less specialized in innovation when comparing with the current equilibrium. In another words, the rise of China leads to the decrease in production cost in the ROW, which makes it possible for each state in US to be more concentrated in innovation and reallocate some of its productions to the ROW. Michigan experiences the largest decline in manufacturing, followed by Massachusetts, Connecticut, and Texas due to the rise of China.[28]

Figure 1.4 plots the percentage changes in innovation productivity against the changes in production workers. When there are no trade shocks, states with a larger increase in production workers will also experience a more substantial growth in innovation efficiency.

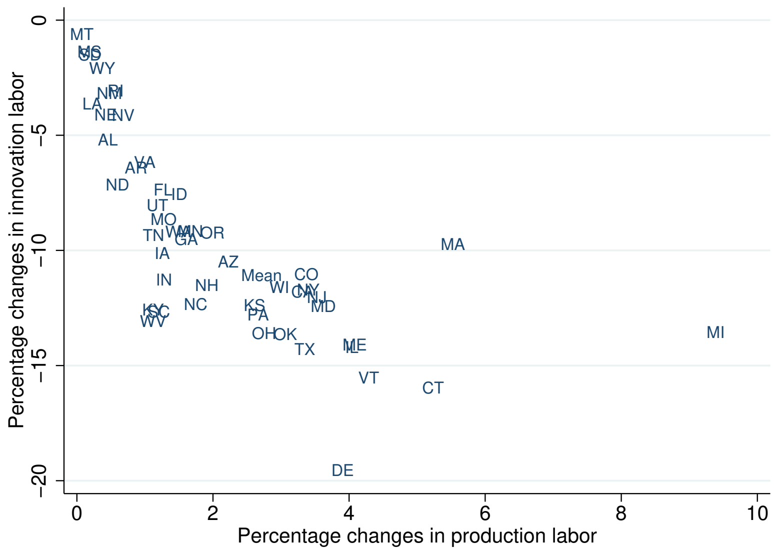
Spillovers versus No Spillovers

I consider another counterfactual exercise by shutting down the spillovers from the local manufacturing to innovation. For cases with and without spillovers from the local manufacturing to innovation, I compute the changes in welfare, innovation productivity, and the number of varieties due to the exogenous trade shocks.

In the following, I will compare the difference in changes across the two scenarios.

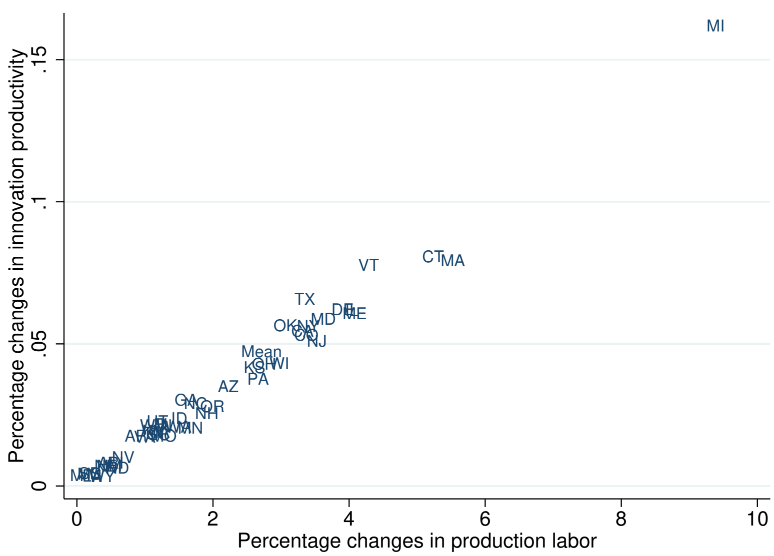
[28] Autor et al. (2016) also find that these states face the highest exposure to trade shocks.

Figure 1.3 Reallocation of Innovation and Production Labor



Notes: This figure plots the percentage changes in innovation and production labor for each state when removing the import competition from China.

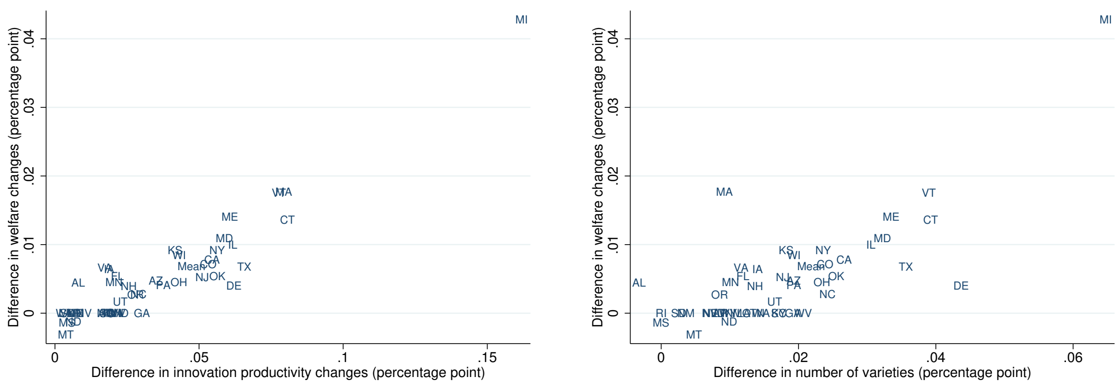
Figure 1.4 The Impact of Production Labor Reallocation on Innovation Efficiency



Notes: This figure plots the percentage changes in innovation productivity against the changes in production workers when removing the import competition from China.

In Figure 1.5, the y-axis in both subfigures plots the difference in welfare changes. The x-axis in the left panel reports the difference in innovation productivity. When there are positive spillovers from the local manufacturing to innovation, the increase in production employment leads to an increase in innovation efficiency. With higher productivity in innovation, it obtains higher welfare compared with the case of no spillovers. The scatterplot of the left panel indicates that the increase in innovation productivity due to the positive spillovers from the local manufacturing results in higher welfare. More than that, higher productivity in innovation also leads to a larger investment in R&D. As a result, more workers will be allocated to innovation activities. The increase in innovation labor further results in the increase in the number of varieties being created, as plotted on the x-axis in the right panel of the Figure 1.5.

Figure 1.5 Difference in Welfare Changes



(a) Productivity in innovation

(b) Number of varieties

Notes: This figure compares the difference in welfare changes, innovation productivity, and the number of varieties for cases with and without spillovers from local manufacturing to innovation. The y-axis in both subfigures plots the difference in welfare changes. The x-axis in the left panel reports the difference in innovation productivity and in the right panel displays the difference in the number of varieties.

1.5 Conclusion

This paper finds evidence that production proximity is crucial for innovation efficiency. I document two novel patterns on the spatial distribution of innovation and production: (i) innovation activities are more agglomerated than production ones, and (ii) innovation and production activities are geographically concentrated. Motivated by the stylized facts, I propose that local manufacturing can enhance innovation and develop an empirical model to allow for the spillovers from the local manufacturing to innovation. In my empirical model, the increment in plant performance depends both on its investment in R&D, and the interaction between R&D and local manufacturing workers. I estimate the model by using a unique confidential plant-level panel data on innovation and production from the US Census Bureau. My estimates show that with more manufacturing workers in the local area, the returns to R&D are higher. My empirical finding is consistent with the idea that geographic proximity facilitates the transmission of knowledge (Audretsch and Feldman, 2004). I extend Arkolakis et al. (2018) to incorporate the positive spillovers from production to innovation that I find in the empirical study, calibrating it to 48 states in the US and the ROW. I then evaluate the impact of China's trade shock on innovation efficiency through the new channel of local spillovers. I find that states with a more significant decline in the manufacturing sector due to the China shock experience a more substantial loss in innovation efficiency.

Chapter 2

Price Stickiness Along the Income Distribution and the Effects of Monetary Policy[1]

Co-authored with Javier Cravino and Andrei A. Levchenko

2.1 Introduction

There is growing recognition that monetary policy shocks have distributional consequences. An active literature argues that monetary policy can have differential effects across various types of agents: savers vs. borrowers (Doepke and Schneider, 2006), financially constrained vs. unconstrained (Williamson, 2008), or young vs. old (Wong, 2016). In turn, the heterogeneity in the impact of monetary policy across agents can determine its overall effectiveness (Auclert, 2017; Beraja et al., 2017; Kaplan et al., 2018). Coibion et al. (2017) show empirically that monetary contractions increase both income and consumption inequality. In all of these contributions, the distributional consequences of monetary policy arise from its heterogeneous impact on the value of agents' income or wealth.

This paper proposes and quantifies a novel mechanism through which monetary policy shocks have distributional consequences. If the effects of monetary shocks on prices are heterogeneous across types of goods (Boivin et al., 2009), and consumption baskets differ across the income distribution (e.g., Almås, 2012), then shocks will differentially affect the prices faced by households of different incomes. We document

[1]This chapter is published in the *Journal of Monetary Economics* (2018). We are grateful to the editor (Yuriy Gorodnichenko), an anonymous referee, Andres Blanco, John Leahy, Indrajit Mitra, and workshop participants at different institutions for helpful comments, to Geoffrey Paulin, David Johnson, Steve Henderson, and John Bieler for help with the CES data, and to Sam Haltenhof for excellent research assistance. Financial support from the National Science Foundation under grant SES-1628879 and from the Washington Center for Equitable Growth is gratefully acknowledged.
Email: jcravino@umich.edu, tinglan@umich.edu, alev@umich.edu.

that the prices of the goods consumed by high-income households are (i) more sticky and (ii) less volatile than those of the goods consumed by middle-income households. We then use both econometric estimates and a New Keynesian DSGE model to quantify the distributional consequences of monetary policy shocks. Both methodologies indicate that these consequences are large relative to the aggregate impact of monetary policy on prices.

Our analysis uses three main sources of data. The first is the US Consumer Expenditure Survey (CES), from which we obtain expenditure shares across detailed product categories for households at different percentiles of the income distribution. The second is the item-level consumer price data from the BLS, which are the most finely disaggregated consumer prices publicly available for the US. Finally, we employ the measures of price stickiness constructed by Nakamura and Steinsson (2008), who report the frequency of price adjustment (i.e. the probability that a price changes in a particular month) for every detailed product category in the US CPI.

We combine these data to compute the average frequencies of price changes for the baskets of goods purchased by households at each income percentile in the CES. We find systematic differences in the price-stickiness of the consumption baskets of different households. On average, 22% of the goods consumed by households in the middle of the income distribution change prices in a given month. However, the frequency of price changes is 24% lower for the goods consumed by the richest percentile.[2] We also compute income-specific consumer price indices (CPIs), following the procedure that the BLS adopts for computing the aggregate CPI.[3] We show similar differences in the volatility of prices faced by different households: the standard deviation of the CPI of the top percentile is 38% lower than that of the CPI of the middle-income households.

These differences across consumption baskets imply that income-specific CPIs may respond differentially to monetary policy shocks. In particular, the CPIs of high-income households should be less responsive to monetary shocks than the CPIs in the middle of the income distribution. We evaluate this hypothesis both econometrically and quantitatively. We first estimate the impulse responses of income-specific CPIs to monetary policy shocks identified using the narrative approach of Romer and Romer

[2]These numbers correspond to frequencies of regular price changes (i.e. excluding sales). The results are similar for the frequency of all price changes (including sales).

[3]When building aggregate consumer price indices, the BLS periodically changes the base year for expenditure weights. In computing income-specific CPIs, we follow the BLS procedure for switching base years after 2004. The information on income is less reliable in the CES prior to 2004, and thus we use 2004 household-specific expenditure weights for CPIs prior to 2004. Using official BLS weights or 2004 aggregate weights produces nearly identical pre-2004 aggregate CPI. See Appendix B.1.2 for more detail.

(2004), as extended by Coibion et al. (2017). We compute the impulse responses using the local projections method (Jordà, 2005). Our estimates show that after 36 months, the CPIs of high-income households respond by about one-third less to the same monetary policy shocks than the CPIs of the middle-income households. Thus, the differences in price stickiness and inflation volatility across consumption baskets have the expected impact on the differential responses of households-specific CPIs to monetary policy shocks in the data.

We then perform a quantitative assessment using a multi-sector, multi-household model with Calvo-style nominal rigidities. In the model, sectors are heterogeneous with respect to their price stickiness, and households are heterogeneous with respect to their income levels and consumption baskets. We calibrate the model to the observed levels of price stickiness and observed cross-household differences in consumption patterns, and simulate the model's response to a monetary policy shock, paying special attention to how a monetary shock differentially affects households. As expected, high-income households' CPIs respond less to a monetary policy shock than middle-income households' CPIs. The difference is once again quantitatively large: after 12 months, the CPI of the households in the top percentile of the income distribution responds by 13% less than that of the middle-income households. We also show that shifting the distribution of income towards households that consume more sticky goods (i.e. more income inequality) would increase the effectiveness of monetary policy, although this effect is modest for realistic changes in inequality.

Our paper draws on, and contributes to, two literatures. The first is the research agenda on the distributional aspects of monetary policy reviewed above. The second is the literature on the differential responses of prices faced by different consumers following macroeconomic shocks. Cravino and Levchenko (2017) document that after a large devaluation in Mexico, consumption price indices of high-income households increased by far less than consumption price indices of the poor. Argente and Lee (2015) show that in the US Great Recession, prices of groceries and general merchandise items consumed by the poorer households increased by more than those consumed by the richer households, while Jaravel (2017) shows that over the past 15 years, product variety increased the most, and inflation was lowest, for the consumption basket of the high-income households. Kaplan and Schulhofer-Wohl (2017) document substantial cross-sectional dispersion in household inflation rates, while Coibion et al. (2015) study the impact of local economic conditions on the geographical variation in prices paid by consumers. Kim (2018) shows that low-quality brands change prices more frequently than high-quality brands within narrow product categories, and evaluates the impact of monetary policy across consumers buying goods of different qualities. Ongoing work by Clayton et al. (2018) focuses on differences in price stickiness of goods consumed by, and produced by, college-educated workers. Our paper documents new facts and proposes a novel mechanism that is based on differential price stickiness of consumption items along the income distribution.

The rest of the paper is organized as follows. Section 2.2 lays out a simple model that illustrates the main mechanism at work, and highlights the key objects of interest that should be the focus of the empirical analysis. Section 2.3 describes the data and documents consumption basket differences across households. Section 2.4 presents the econometric evidence, and Section 2.5 presents the quantitative model and reports the responses of household-specific inflation to an aggregate monetary shock. Section 2.6 concludes.

2.2 A Simple Sticky Price Model

Before presenting our data, we describe a simplified sticky price model to build intuition on how aggregate shocks can have distributional consequences when nominal rigidities are heterogeneous across goods and households consume different baskets of goods.

Setup: Consider a two-period economy populated by H types of households indexed by h, each consuming a different basket of goods. In the first period, the state of the world is known, and in the second period the economy can experience one of infinitely many shocks or states, s.[4] The (log) price of the consumption basket (i.e. the CPI) consumed by household h in period t is given by

$$p_t^h(s) \equiv \sum_j \omega_j^h p_{j,t}(s),$$

where ω_j^h is the share of goods from sector j in household h's consumption basket. We define the aggregate price index as $p_t(s) \equiv \sum_h s^h p_t^h(s) = \sum_j \omega_j p_{j,t}(s)$, where s^h denotes household h's share in the aggregate consumption expenditures, and $\omega_j \equiv \sum_h s^h \omega_j^h$ is the economy-wide expenditure share in sector j.

Sectoral goods are aggregates of a continuum of intermediates that are produced by monopolistically competitive firms. We introduce price stickiness by assuming

[4]The set of shocks can include monetary shocks, but at this stage we do not need to specify the exact nature of the shocks.

that in the second period, only a fraction θ_j of producers in each sector j can observe the realization of the state before setting their prices. The remaining producers must set prices before observing the realization of the state. To isolate the role of sectoral differences in price rigidities, we assume all producers operate the same CRS technology and set constant markups. In the first period, all the producers know the state and so they set the same price, which we label p_1. In the second period, all producers that observe the state set the same price, which we label $\bar{p}_2(s)$. The producers that don't observe the state set a price that we label p_2^e. Note that p^e is not a function of the state. Without loss of generality we assume that the shocks are mean zero, so that $p_2^e = p_1$.

The average price in sector j in the second period is then given by:

$$p_{j,2}(s) = \theta_j \bar{p}_2(s) + [1 - \theta_j] p_1. \tag{2.1}$$

Let $\pi^h \equiv p_2^h(s) - p_1$ define the household-specific inflation rate. The difference in inflation faced by two households, h and h', is:

$$\pi^h(s) - \pi^{h'}(s) = [\bar{p}_2(s) - p_1] \sum_j \left[\omega_j^h - \omega_j^{h'}\right] \theta_j.$$

This expression highlights that the difference between two households' CPIs is driven by the covariance between the differences in their expenditure shares across sectors, $\omega_j^{h'} - \omega_j^h$, and the price stickiness of those sectors, θ_j. Households that consume less price-sticky goods will experience larger CPI changes following a shock than households consuming relatively more price-sticky goods. Dividing by the aggregate inflation $\pi(s) \equiv p_2(s) - p_1$, yields an expression relating the differences in household-specific inflation to objects that can be measured in the data:

$$\frac{\pi^h(s) - \pi^{h'}(s)}{\pi(s)} = \frac{\bar{\theta}^h - \bar{\theta}^{h'}}{\bar{\theta}}, \tag{2.2}$$

where $\bar{\theta}^h \equiv \sum_j \omega_j^h \theta_j$ and $\bar{\theta} \equiv \sum_h s^h \bar{\theta}^h$. Note that this expression is independent of the realization of the state.

Discussion: Equation (2.2) shows how aggregate shocks can have distributional consequences when price rigidities are heterogeneous across goods and households consume different baskets of goods. In this simple model where all firms face the same costs and markups are constant, the weighted average frequencies of price changes, $\bar{\theta}^h$, are sufficient statistics for all the distributional consequences, irrespective of the nature of the aggregate shocks. Equation (2.2) states that, in response to a shock that generates positive inflation, inflation will be relatively high for households consuming goods with relatively more flexible prices (i.e. high $\bar{\theta}^h$).

To get a sense of the magnitude of these distributional consequences we can do a back of the envelope calculation using US data (described in detail below). In our data, $\bar{\theta}^t \approx 0.17$ for households in the top percentile of the income distribution, $\bar{\theta}^m \approx 0.22$ for households at the middle of the income distribution, and $\bar{\theta} \approx 0.21$. These numbers result in $\frac{\bar{\theta}^t - \bar{\theta}^m}{\bar{\theta}} \approx -0.24$, which indicates that a shock that increases the aggregate CPI relative to its unconditional mean by 1% will also generate a -0.24% gap between the price of the consumption baskets consumed by the top vs. the middle of the income distribution.

The simple model also illustrates the connection between sectoral price stickiness and sectoral price volatility. From (2.1), we can see that sectoral inflation, $\pi_j(s) \equiv p_{j,2}(s) - p_1$, is less volatile in more sticky-priced sectors:

$$\sigma_{\pi_j} = \theta_j \sigma_{\bar{p}},$$

where σ_{π_j} is the standard deviation of inflation in sector j price, and $\sigma_{\bar{p}}$ is the unconditional standard deviation of $\bar{p}_2(s)$. The ratio of standard deviations of sectoral inflation relative to the standard deviation of aggregate inflation is then given by the ratio of the sectoral to the aggregate frequency of price changes:

$$\frac{\sigma_{\pi_j}}{\sigma_\pi} = \frac{\theta_j}{\bar{\theta}}, \tag{2.3}$$

Differences in sectoral price volatility translate into differences in household-level CPI volatility. The standard deviation of household-specific inflation, normalized relative to the aggregate is:

$$\frac{\sigma_{\pi^h}}{\sigma_\pi} = \frac{\bar{\theta}^h}{\bar{\theta}}. \tag{2.4}$$

Households consuming more price-sticky goods experience less volatile price changes. The following section evaluates the relationships (2.3) and (2.4) in the data. Of course, these relationships may not hold if the standard deviation of the desired price change $\sigma_{\bar{p}}$ is sector-specific (as would be the case for example if there are sector-specific shocks).

To summarize, our illustrative model establishes that in order to understand how the CPIs of different households react to monetary or other shocks, we must examine the differences in price stickiness of consumption baskets across households. In addition, it suggests a one-to-one relationship between sectoral price stickiness and sectoral price volatility. Thus, a closely related object to be examined in the data is differences in inflation volatility across households.

2.3 Empirical Findings

This section describes our data sources and documents our two empirical findings on how consumption baskets differ across the income distribution. It then evaluates the relationship between frequencies of price changes and inflation volatility suggested by equations (2.3) and (2.4). Appendix B.1 describes in detail the construction of expenditure shares from the CES and of the income-specific CPIs.

2.3.1 Data

We combine data on expenditure shares from the CES with the item-level consumer prices from the BLS and with the frequency of price adjustment data from Nakamura and Steinsson (2008). The CES contains two main modules, the Interview and the Diary. The Interview module collects responses from about 30,000 households annually, and asks households about the purchases they make in all categories, as well as other demographic information. Each household is interviewed for up to 4 consecutive quarters in the Interview module. The Diary module surveys about 10,000 households per year, at weekly frequency. The Diary questionnaire contains detailed questions about daily purchases, such as groceries. All in all, there are questions on 350 distinct expenditure categories in the Interview module, and on 250 distinct grocery and related categories in the Diary module.

The large majority of households do not report expenditures in all possible categories in a given year. In addition, a different set of households is surveyed in the Interview and in the Diary files, so the full consumption profile (both Diary and Interview module expenditures together) of any particular household is never observed. This means that we cannot compute expenditure shares for each household. Rather, we aggregate households into percentiles and work with percentile-level expenditure shares. Each percentile contains about 300 households responding to the Interview questions, and 100 households responding to Diary questions. Table B.1 in Appendix B.1.1 reports the income cutoffs and average incomes in the selected quantiles of the income distribution. It is important to note that income categories in the CES (such as wage income) are subject to top-coding. Nonetheless, there is a great deal of variation in incomes of households present in the CES, with incomes of the top 5% of households an order of magnitude higher than those at the median. Throughout the paper, the percentiles of the income distribution are defined based on income information in the CES rather than any external data source.

We use these data to compute the measures of income-specific frequencies of price changes, price indices, and price volatility defined in Section 2.2. The average frequencies of price changes, $\bar{\theta}^h = \sum_j \omega_j^h \theta_j$, combine the income-specific expenditure weights ω_j^h from the CES with the product-specific frequencies of price changes θ_j from Nakamura and Steinsson (2008).[5] To compute them, we match CES expenditure categories to the Entry Level Items (ELIs), a basic category in the CPI for which Nakamura and Steinsson (2008) report frequencies. There are a total of 265 ELI categories. In this exercise, we use the expenditure shares from the year 2015, but the results are quite similar for expenditure shares in other years.

We calculate household-specific inflation as $\pi_t^h \equiv p_t^h - p_{t-1}^h$, where the household-specific price indices are given by $p_t^h \equiv \sum_j \omega_{j,\tau}^h p_{j,t}$. The time-varying income-specific expenditure weights $\omega_{j,\tau}^h$ come from the CES, and are updated following the procedure used by the BLS to compute the aggregate CPI. See Appendix B.1.2 for the complete description of the procedure.[6] The item-level price indices $p_{j,t}$ also come from the BLS. The item level is the finest publicly available level of disaggregation in the US CPI data (the BLS does not report inflation numbers at the ELI level), and is slightly more coarse than ELI, containing 178 distinct expenditure categories starting in 1998. The price data are monthly, for the period 1969-2008, though prior to 1998 the BLS used a different product classification. We take 12-month log-differences to obtain annual growth rates. We then compute the standard deviations of those annual growth rates for the price indices at each income level.

2.3.2 Two Facts About Consumption Basket Differences Across Households

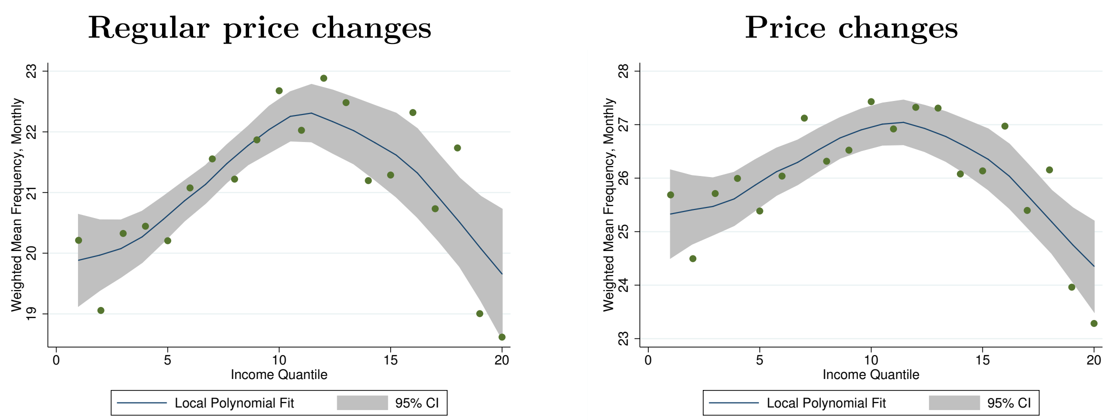
Fact 1: Prices of goods consumed by middle-income households are relatively flexible. Figure 2.1 presents the scatterplot of the weighted mean frequency

[5] Nakamura and Steinsson (2008) calculate these frequencies as the fraction of prices that change in a given a month, both for all prices, and for regular prices (excluding sales).

[6] Like aggregate measures of inflation, our household-specific price indices are subject to substitution bias. This bias is second-order, and is likely to be negligible for realistic monetary policy shocks. Appendix B.3 shows that the differences in inflation rates computed from a Laspeyres vs. a Paasche price index between 1987 and 2016 are an order of magnitude smaller than the inflation rate.

of price adjustment, $\bar{\theta}^h$, for households at each of the 20 quantiles of the income distribution in the CES. Thus, each dot corresponds to 5% of households. The solid line through the data is the local polynomial fit, and the shaded area is the 95% confidence interval. The left panel depicts $\bar{\theta}^h$ when θ_j is measured as the frequency of regular (non-sale) price changes, while the right panel measures θ_j as the frequency of all price changes, including sales. Mean frequencies of price changes are hump-shaped along the income distribution: middle-income households consume goods with more frequent price changes, while high- and low-income households consume goods with less frequent price changes. Table 2.1 summarizes the underlying magnitudes. It

Figure 2.1 Weighted mean frequency of price changes



Notes: This figure plots the weighted mean frequency of price changes for households in 20 quantiles of the income distribution. Each dot represents 5% of the income distribution.

reports, for different slices of the income distribution, the weighted mean frequency of price adjustment. For the households around the median – the 40-60 income percentiles – the frequency of regular price adjustment is 22.16 percent per month. For all the households from the 1st to the 95th percentile, that frequency is 21.17 percent per month. By contrast, the frequency falls to 19.27 for the households in the 96th to 99th percentile, and further to 16.82 for the top percentile in the distribution. Thus, the weighted mean frequency of price adjustment is some 25% lower for the households in the top 1% of income compared to the households around the median income. Including sales, the results are quite similar. In particular, the top 1% of the income distribution has an 18% lower weighted mean frequency of price adjustment than the middle of the income distribution.

Table 2.1 Weighted mean frequency of price changes and CPI volatility at different points on the income distribution

	Income percentile			
	40-60	1-95	96-99	100
Frequency of price changes:				
Regular prices	22.16	21.17	19.27	16.82
All prices (incl. sales)	26.90	26.16	23.75	22.17
Standard deviation of CPI:	0.021	0.020	0.015	0.013

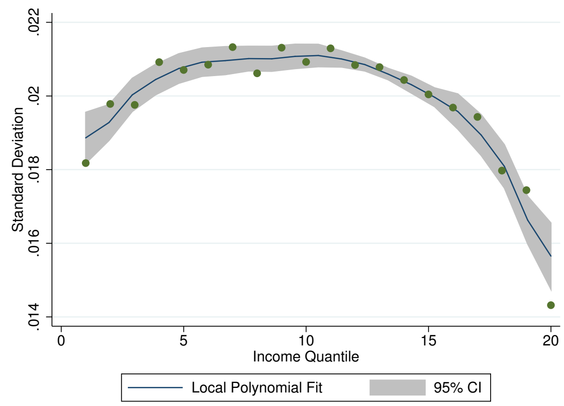
Note: This table reports the weighted mean frequency of price changes, and the standard deviation of the 12-month log change in CPI for consumers of different incomes.

Fact 2: Prices of goods consumed by middle-income households are relatively volatile. Figure 2.2 reports the standard deviation of π_t^h, the income-specific inflation. Inflation volatility is also hump-shaped along the income distribution. The households with middle incomes experience the highest inflation volatility, whereas the lowest volatility is found at the top of the income distribution. The bottom of Table 2.1 reports the values of the standard deviation of inflation faced by consumers of different incomes. The annual inflation rate has a standard deviation of 0.020 for consumers in the bottom 95% of the income distribution, and 0.021 for consumers in the middle (40-60th percentiles). By contrast, the standard deviation of annual inflation is 0.015 for households in the 96th to 99th percentile of the income distribution, and 0.013 for those in the top 1%.

Discussion: What consumption patterns are responsible for these differences in price stickiness and volatility across baskets? Table B.3 reports the 10 consumption items with the largest differences in the expenditure shares between the middle 20% of the income distribution and the top 1% of the income distribution.

The top categories in which the middle-income consumers exhibit highest expenditure shares relative to the top 1% are mainly goods such as Gasoline, Electricity, Motor Vehicle Insurance, and Used Cars. The items with the largest expenditure shares of the top 1% relative to the middle income consumers are mostly services, such as Elementary School and College Tuition, Child Care, Airfare, Domestic Services,

Figure 2.2 Standard deviation of the changes in consumption price indices



Notes: This figure plots the standard deviation of the 12-month log difference in the consumption price indices for households in 20 quantiles of the income distribution. Each dot represents 5% of the income distribution.

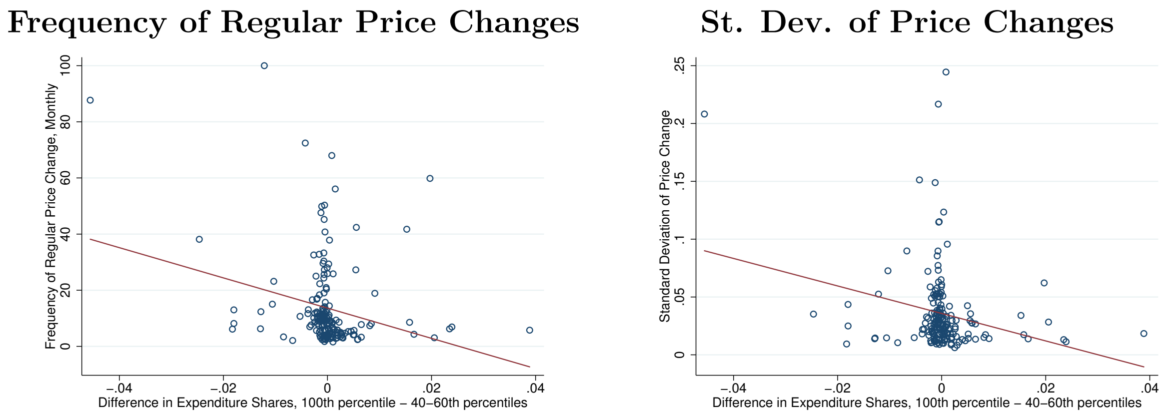
and Club Membership Fees. Among the 10 categories consumed more intensively at the middle of the income distribution, the frequency of monthly price adjustment is in excess of 30%. Among the 10 items most disproportionately consumed by the top 1%, the frequency of regular price adjustment is 16%, and total price adjustment 18%. In either case, the difference in average price adjustment frequency between these two sets of items is pronounced. The pattern of price stickiness is not universal. Among the top 1%'s (relative) top 10 items is Airfare, with price adjustment frequency of almost 60% per month. On the flip side, General Medical Practice and Limited Service Meals are in the middle 20%'s top 10, and among the price-stickiest categories.

The left panel of Figure 2.3 plots the frequency of the regular price adjustment on the y-axis against the difference in the expenditure shares between the top 1% and the middle 20%, with positive values meaning that the top 1% has higher expenditure shares in that category. The majority of categories are concentrated on 0, implying that the high- and the middle-income categories have similar expenditure shares. There is a large range, however, and all in all the relationship between these relative shares and the frequency of price adjustment is negative. The correlation between the x-axis and y-axis variables is -0.251.

The categories with the largest expenditure share differences also differ substantially in the standard deviation of item-level price changes. The mean standard deviation of 12-month log price changes in the set of goods consumed most disproportionately by the middle-income households is 0.049, more than double the 0.023 mean in the set of goods consumed by high-income households.

The outlier sector here is Gasoline, whose standard deviation is 0.208, and which

Figure 2.3 Expenditure differences, frequency of price changes, and standard deviation of price changes



Notes: The left panel plots the frequency of price changes against the difference in sectoral expenditure shares between households in the top 1% and the middle 20% of the income distribution. The right panel plots the standard deviation of 12-month log price change against the difference in sectoral expenditure shares between households in the top 1% and the middle 20% of the income distribution. Both panels include the OLS fit through the data.

is also the sector with the single largest expenditure share discrepancy – in either direction – between the middle- and high-income households. But the differences persist even if we focus on the median standard deviation, or drop Gasoline when computing the mean.[7] The right panel of Figure 2.3 displays the scatterplot of the standard deviation of log price change at the item level against the expenditure share difference between the high- and middle-income consumers. Once again, most expenditure share differences are close to zero. Nonetheless, the correlation between the expenditure share differences and standard deviation of price changes is negative at -0.255.

2.3.3 Frequency of Price Changes and Inflation Volatility

This section evaluates the relationship between frequency of price changes and inflation volatility suggested by equations (2.3) and (2.4) of Section 2.2, by providing the data counterparts of those postulated relationships. The left panel of Figure 2.4 plots the empirical counterpart of (2.3), along with a 45-degree line. As (2.3) expresses

[7] If we exclude Gasoline, the differences across households reported in Figure 2.1 and Table 2.1 are attenuated, but the basic patterns hold. Gasoline appears to be responsible for about half of the difference in the weighted average frequency of price adjustment between the top-income and the middle-income households. Dropping Gasoline, Figure 2.2 is somewhat modified. It is still true that high-income households have lower inflation volatility than middle-income ones, but now the highest inflation volatility occurs in the bottom income tercile.

both the right- and left-hand side variables relative to the average, we rescale both the product-level standard deviation and the frequency of price adjustment by their means across items. Each dot represents one of the 178 disaggregated CPI items. A positive relationship with a slope close to unity is evident in this plot; the correlation coefficient between these two variables is 0.715.

The right panel plots the empirical counterpart of (2.4), once again with both y- and x-axis variables rescaled by their respective means and adding a 45-degree line. Each dot represents 5% of the income distribution, as in Figures 2.1-2.2. There is an evident positive relationship between these two variables, with the correlation coefficient of 0.643. Households consuming more flexible-priced goods tend to experience higher CPI volatility. This is not surprising, as we are in effect plotting the y-axes of Figures 2.1 and 2.2 against each other, and both follow a similar inverse U-shape with respect to income quantile.

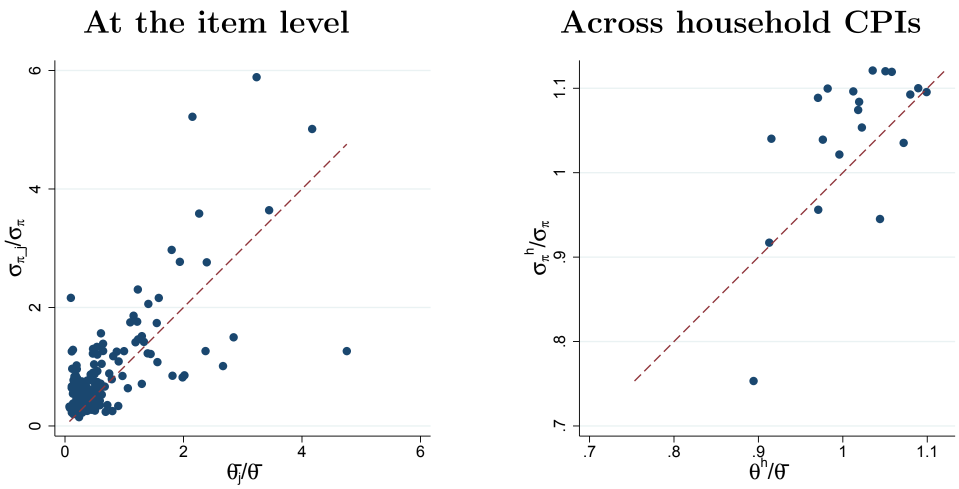
Figure 2.4 is consistent with the prediction of our time-dependent pricing model from Section 2.2 that more sticky sectors should have less volatile prices. If the frequencies of price adjustment are endogenous, it could be that the differences in the volatility of sectoral shocks are what drives the frequency of price adjustment. Note that irrespective of the direction of causality, the correlation between inflation volatility and frequency of price adjustment across households of different incomes is that depicted in Figure 2.4. The following section shows that some households are more sensitive to monetary policy shocks than others. For those results, we do not need to specify whether the difference in the frequency of price adjustment across sectors is exogenous or driven by the volatility of sectoral shocks.[8]

2.4 Impulse Responses of Income-specific CPI to Monetary Policy Shocks

The previous section shows that prices of goods consumed by high-income households are more sticky and less volatile than those of the goods consumed by middle-income households. This suggests that monetary shocks can have distributional consequences by affecting the relative prices of consumption baskets of households at different points on the income distribution. We now present evidence that monetary policy shocks indeed lead to smaller CPI changes for households at the top of the income distribution

[8] Indeed, Boivin et al. (2009) provide evidence that prices of more volatile goods react systematically more strongly to monetary policy shocks.

Figure 2.4 Stickiness and volatility



Notes: The left panel plots the standard deviation of 12-month log price change at the item level vs. the frequency of price adjustment for that item. The right panel plots the standard deviation in the 12-month log change in overall household CPI against the weighed mean frequency of price adjustment for that household type; each dot represents 5% of the income distribution. Both plots include the 45-degree line.

relative to the middle. Our baseline specification adopts the local projection method of Jordà (2005) to estimate the responses of income-specific CPIs to monetary policy shocks. Online Appendix B presents impulse responses of income-specific price indices using the FAVAR methodology following Bernanke et al. (2005) and Boivin et al. (2009).

The local projection method estimates regressions of the dependent variable at horizon $t+s$ on the shock in period t and uses the coefficient on the shock as the impulse response estimate. We estimate the following series of regressions:

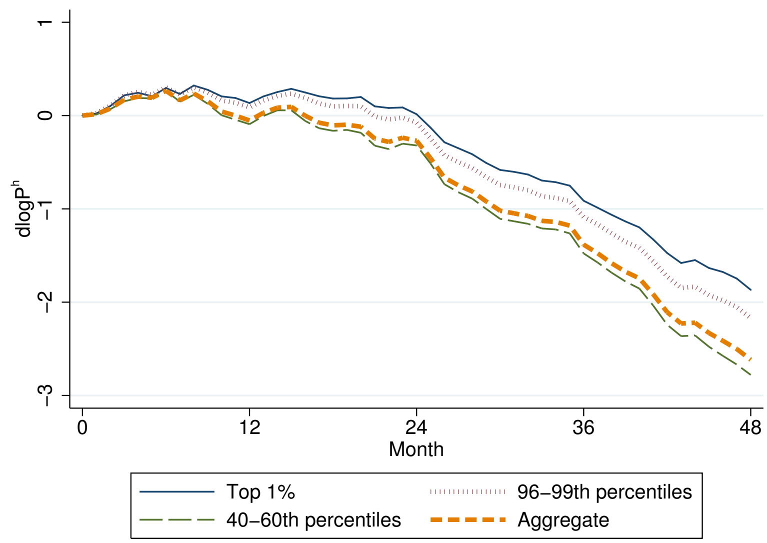
$$p^h_{t+s} - p^h_t = \alpha_s + \theta_s shock^{RR}_t + \sum_{j=1}^{J} \beta_{s,j}(p^h_{t+1-j} - p^h_{t-j}) + \sum_{i=1}^{I} \gamma_{s,i} shock^{RR}_{t-i} + \epsilon_{t+s}. \quad (2.5)$$

Here, p^h_t is the log of income-specific CPIs, and $shock^{RR}$ is the Romer and Romer (2004) narrative-based measure of monetary policy shocks from Coibion et al. (2017). The control variables include 48 lags of the shocks ($I = 48$) and 6 lags of monthly income-specific inflation ($J = 6$). The coefficient θ_s gives the response of income-specific prices at $t+s$ to a monetary policy shock at t. We estimate impulse responses over a horizon of 48 month with $s = 0, 1, \ldots, 48$.

In our application, we estimate the impulse response of income-specific prices for each income percentile. We use monthly data for the sample period 1969m1 to

2008m12. Figure 2.5 plots the estimated impulse responses of income-specific prices for selected percentiles to a 100-basis-point of contractionary monetary policy shock. The consumer price indices of the high-income households react substantially less to monetary policy shocks than those for the middle of the income distribution. The difference is economically meaningful. After 36 months, the top-1% households' CPI responds by 38% less, and the 96-99th percentile households by 26% less, than the CPI of the households in the middle of the income distribution (40-60th percentiles). After 48 months, the differences are still 33% and 22%, respectively.

Figure 2.5 Income-specific CPI impulse responses to a monetary policy shock

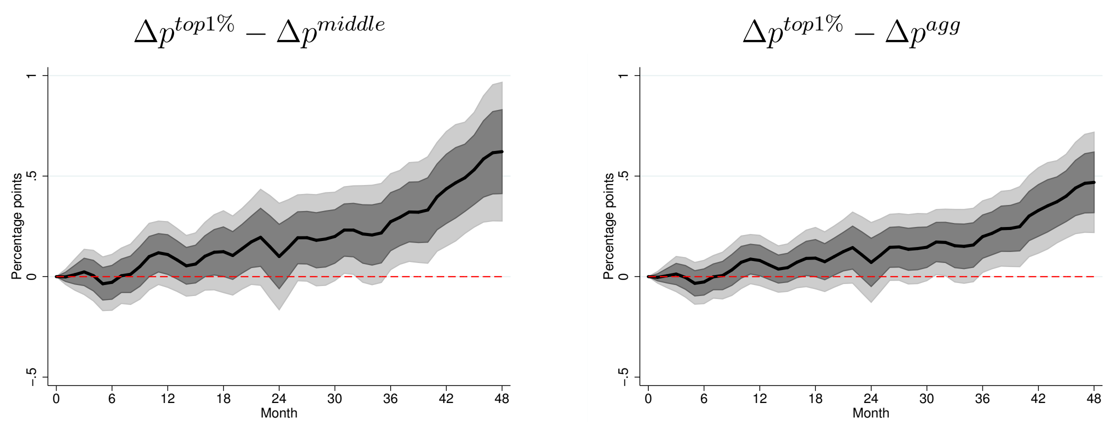


Notes: This figure plots the impulse responses of income-specific price indices to a monetary policy shock identified using the narrative approach of Romer and Romer (2004), as extended by Coibion et al. (2017). The impulse responses are computed using the local projections method (Jordà, 2005)

Our main object of interest is not the overall response of prices to a monetary shock, but rather the differential response of the CPIs of different households. We estimate a version of equation (2.5) using the difference between the (log) CPI of the top 1% and the (log) CPI of the middle 20% of the income distribution, and the difference between the top 1% and the aggregate CPI to quantify the effect of monetary policy on inflation faced by different households. Figure 2.6 plots the difference in the response of the CPIs of different households. The dark and light grey areas indicate 1 and 1.65 standard deviation confidence intervals, respectively. The figures show that following a contractionary monetary shock, the price level for the the top income households falls by less than the price level for middle income households (so that the difference between the two is positive). The difference is statistically significant, and is about a third of the size of the response of aggregate inflation to the same monetary shock reported in Figure 2.5.

The econometric evidence thus suggests that monetary shocks can have large distributional effects across households of different incomes. The following section complements this evidence using a New Keynesian model that quantifies the mechanisms described in Section 2.2.

Figure 2.6 Differences in inflation changes between income groups



2.5 Quantitative Framework

This section sets up a sticky price model with multiple households and sectors to evaluate how monetary shocks affect consumption price indices for households at different points of the income distribution.

2.5.1 Setup

Preliminaries: We consider an economy populated by H types of households indexed by h. Households get utility from consuming a bundle of goods produced by J different sectors of the economy indexed by j. Sectoral goods are produced by aggregating the output of a continuum of monopolistic intermediate producers indexed by i. The monetary authority sets the nominal interest rate following a Taylor rule.

Households: Each type of household h has preferences given by:

$$U^h = \mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[lnC_t^h - N_t^h \right], \tag{2.6}$$

and faces the budget constraint:

$$P_t^h C_t^h + \Theta_{t,t+1} B_{t+1}^h = W_t A^h N_t^h + T_t^h + B_t^h. \tag{2.7}$$

Here, C_t^h is the bundle of goods consumed by households of type h, and P_t^h is the price of this bundle. N_t^h and A^h respectively denote labor supply and the efficiency of household h, and W_t is the nominal wage per efficiency unit. B_{t+1}^h is a bond that pays one dollar in $t+1$, and $\Theta_{t,t+1}$ is the date t price of that bond. Finally, T_t^h are transfers to the households from the government and from firms' profits.

The bundle of goods consumed by each type of household is:

$$C_t^h = \left[\sum_j^J \left[\overline{\omega}_j^h \right]^{\frac{1}{\eta}} \left[C_{j,t}^h \right]^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}}, \tag{2.8}$$

where $C_{j,t}^h$ denotes household h's consumption of final goods from sector j, and $\overline{\omega}_j^h$ is a household-specific taste shifter for sector j. Note that the parameters $\overline{\omega}_j^h$ are associated with a particular household h that has efficiency A^h. These parameters allow us to capture in a reduced form the non-homotheticicities that may lead to the cross-household differences in expenditure shares observed in the data. The model will be calibrated directly to household-specific expenditure shares. The price index associated with this bundle is:

$$P_t^h = \left[\sum_j^J \overline{\omega}_j^h P_{j,t}^{1-\eta} \right]^{\frac{1}{1-\eta}},$$

where $P_{j,t}$ is the price of the sector j aggregate. Note that both C_t^h and P_t^h are indexed by h, as the bundle (2.8) differs across households. Monetary shocks can differentially affect households if households put different weights across sectors and shocks have heterogeneous effects across sectoral prices $P_{j,t}$.

Sectoral Demands and Technologies: The demand function associated with the bundle (2.8) is given by:

$$C_{j,t}^h = \overline{\omega}_j^h \left[\frac{P_{j,t}}{P_t^h} \right]^{-\eta} C_t^h.$$

Adding across households, aggregate demand for the final good produced in sector j is

$$P_{j,t}C_{j,t} = \omega_{j,t} \left[\frac{P_{j,t}}{P_t}\right]^{1-\eta} P_t C_t, \tag{2.9}$$

where $P_t C_t$ are aggregate nominal expenditures, $\omega_{j,t} \equiv \sum_h \overline{\omega}_j^h s^h \frac{\left[P_t^h\right]^{\eta-1}}{\sum_h s_h \left[P_t^h\right]^{\eta-1}}$, and $P_t \equiv \left[\sum_j \omega_{j,t} P_{j,t}^{1-\eta}\right]^{\frac{1}{1-\eta}}$. In these expressions, s^h is the share of household h in aggregate expenditures.

Sectoral goods are produced by aggregating the output of a continuum of intermediate producers according to

$$Y_{j,t} = \left[\int Y_{j,t}(i)^{\frac{\gamma-1}{\gamma}} di\right]^{\frac{\gamma}{\gamma-1}}.$$

Total demand faced by intermediate producer i is then:

$$Y_{j,t}(i) = \left[\frac{P_{j,t}(i)}{P_{j,t}}\right]^{-\gamma} Y_{j,t}. \tag{2.10}$$

Intermediate Good Producers: Intermediate producers behave as monopolistic competitors and set prices as in Calvo (1983). The probability that a producer can change its price in any period depends on the sector in which it operates, and is given by θ_j. The producers operate a linear technology

$$Y_{j,t}(i) = \bar{N}_{j,t}(i), \tag{2.11}$$

where $\bar{N}_{j,t}(i)$ denotes the efficiency units of labor used by producer i. The profit-maximizing price for an intermediate producer that gets to adjust prices satisfies:

$$\bar{P}_{j,t} = \arg\max \left\{ \sum_{k=0}^{\infty} [1-\theta_j]^k \mathbb{E}_t \left\{ \Theta_{t,t+k} \left[\bar{P}_{j,t} - W_{t+k}\right] Y_{j,t+k}(i) \right\} \right\} \tag{2.12}$$
$$\text{subject to} \quad (2.10).$$

Monetary Policy: The monetary authority sets nominal interest rates according to a Taylor rule:

$$exp(i_t) = exp(\rho_i i_{t-1}) \left[\Pi_t^{\phi_\pi} \left[\frac{Y_t}{\bar{Y}}\right]^{\phi_y}\right]^{1-\rho_i} exp(\nu_t),$$

where $i_t \equiv -logQ_{t,t+1}$ is the nominal interest rate, $\Pi_t \equiv P_t/P_{t-1}$ is aggregate inflation, and $\bar{Y}$ is the efficient level of output. Finally, ν_t is a monetary shock that satisfies

$$\nu_t = \rho_\nu \nu_t + \varepsilon_{\nu,t}, \tag{2.13}$$

with $\varepsilon_{\nu,t} \sim N\left(0, \sigma_{\varepsilon_\nu}\right).$

Equilibrium: An equilibrium for this economy is a set of allocations for the households $\{C_t^h, C_{j,t}^h, N_t^h\}_{\forall j,h,t}$, sectoral good producers $\{Y_t^j, \left\{Y_t^j(i)\right\}_i, \left\{\bar{N}_t^j(i)\right\}_i\}_{\forall j,t}$, and price policy functions for intermediate producers $\left\{\bar{P}_{j,t}\right\}_{\forall j,t}$, such that given prices: (i) households maximize (2.6) subject to (2.7); (ii) sector j final producers minimize costs according to equations (2.9) and (2.10); (iii) intermediate producers maximize profits by solving (2.12); and (iv) goods and labor markets clear, $\sum_h C_{j,t}^h = Y_t^j$ and $\sum_h A^h N_t^h = \sum_j A^h \int \bar{N}_{j,t}(i)\, di$.

We now characterize the equilibrium of a log-linearized version of this economy around a non-stochastic steady state, following the tradition in the New Keynesian literature. In what follows, we use lower-case letters to denote the log-deviations of a variable from its non-stochastic steady state. The optimality conditions associated with the household problem are the labor-leisure condition:

$$P_t^h C_t^h = A^h W_t,$$

and the Euler equation:

$$\Theta_{t,t+1} = \beta \mathbb{E}_t \left\{ \frac{P_t^h C_t^h}{P_{t+1}^h C_{t+1}^h} \right\}.$$

Adding the labor-leisure condition across households we obtain that each type of household gets a constant share of nominal consumption expenditures, $s^h \equiv \frac{P_t^h C_t^h}{P_t C_t} = \frac{A^h}{A}$, where $A \equiv \sum_h A^h$. Substituting into the optimality conditions and log-linearizing we obtain:

$$w_t - p_t = c_t, \tag{2.14}$$

and

$$c_t = \mathbb{E}_t\left\{c_{t+1}\right\} - \left[i_t - \mathbb{E}_t\left\{\pi_{t+1}\right\} - \rho\right], \tag{2.15}$$

with $\rho \equiv -log\beta$. Goods market clearing implies $y_t = c_t$. Substituting into equation

(2.15) we obtain:

$$y_t = \mathbb{E}_t \{y_{t+1}\} - [i_t - \mathbb{E}_t \{\pi_{t+1}\} - \rho] . \tag{2.16}$$

The optimal log-price that solves (2.12) can be written recursively as:

$$\bar{p}_{j,t} = [1 - \beta [1 - \theta_j]] w_t + \beta [1 - \theta_j] \mathbb{E}_t [\bar{p}_{j,t+1}] ,$$

and the law of motion for the sectoral price indices is

$$p_{j,t} = \theta_j \bar{p}_{j,t} + [1 - \theta_j] p_{j,t-1}.$$

Combining we these two equations we obtain a sectoral Phillips curve,

$$\pi_{j,t} = \lambda_j [w_t - p_{j,t}] + \beta \mathbb{E}_t \{\pi_{j,t+1}\} , \tag{2.17}$$

with $\lambda_j \equiv \frac{\theta_j [1-\beta(1-\theta_j)]}{[1-\theta_j]}$. Finally, the Taylor rule is:

$$i_t = \rho_i i_{t-1} + [1 - \rho_i] [\rho + \phi_\pi \pi_t + \phi_y \tilde{y}_t] + \nu_t. \tag{2.18}$$

Equations (2.14)-(2.18) can be used to solve for all sectoral inflation rates, along with the output gap, real marginal costs, real wages, the nominal interest rate, and the aggregate inflation rate. Sectoral inflation rates can then be used to compute household-specific inflation according to:

$$\pi_t^h = \sum_j \omega_j^h \pi_{j,t}.$$

Note that as a result of taking the first-order approximation, we dropped the time subscripts on the expenditure weights ω_j^h, since changes in prices only affect expenditure shares to a second order. This second-order substitution bias is likely to be negligible for realistic monetary policy shocks, as discussed in Online Appendix C.

In what follows, we will use the model to ask two questions: (i) what is the effect of a monetary policy shock $\varepsilon_{\nu,t}$ on household-specific inflation?, and (ii) how do changes in the distribution of income s^h affect the response of inflation π_t and the output y_t to a monetary shock?

2.5.2 Results

Calibration

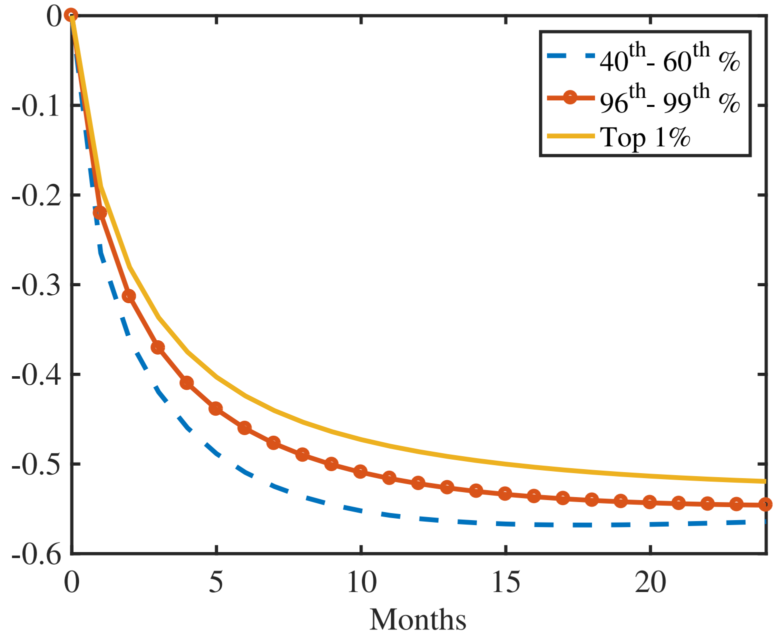
To evaluate the impact of monetary shocks, we need to assign values for the discount factor β, the coefficients in the Taylor rule, ρ_i, ϕ_π and ϕ_y, the process for the shocks, ρ_ν and σ_{ε_ν}, the sectoral frequencies of price changes, θ_j, $\forall_j$, the sectoral household-specific expenditure shares, ω_j^h, and the household shares in aggregate consumption spending, s^h. We calibrate the model to monthly data and use values for most of these parameters that are standard in the literature. In particular, we set $\beta = 0.96^{1/12}$, which corresponds to an annualized real interest rate of 4 percent, and take the Taylor rule parameters $\rho_i = 0.95$, $\phi_\pi = 1.5$ and $\phi_y = 0.5/12$ and set the persistence of the shocks to $\rho_\nu = 0$, as in Christiano et al. (2010). Finally, we calibrate the model to 265 sectors and 20 household types, and calibrate the frequencies of price changes θ_j and the expenditure shares ω_j^h and s^h using the data from Nakamura and Steinsson (2008) and the CES data presented in Section 2.3.

Distributional Consequences of Monetary Shocks

We now evaluate the distributional consequences of a monetary shock in this model. Figure 2.7 plots the impulse response of the household-specific price indices to a one standard deviation shock to $\varepsilon_{\nu.t}$. The figure shows that the shock has distributional effects: prices of the middle-income households are the most sensitive to the shock, and prices are the least sensitive for the top-income households. This is not surprising, since in our model, as in the data, households at the top of the income distribution consume the goods that are the most sticky and thus respond more sluggishly to shocks.

Table 2.2 reports the price indices faced by households at different points of the income distribution following the monetary shock, expressed relative to the aggregate price index. The table shows that the cumulative response after 6 months of the prices faced by the top 1 percent is about 13% smaller than that of the aggregate price index, and almost 20% smaller than the response of the prices faced by the households at the middle 5 percent of the income distribution. These differences are quite persistent, the cumulative change in prices faced by the richest 1% is still 10% smaller than that faced by the middle income households 18 months after the shock.

Figure 2.7 Impulse responses of household-specific CPIs to a monetary shock



Notes: This figure plots the impulse responses of income-specific CPIs to a monetary policy shock, simulated using the model in this section.

Table 2.2 Cumulative inflation, relative to aggregate

	Bottom 5%	Middle 5%	96-99 %	Top 1%
6 months	0.993	1.059	0.952	0.874
12 months	1.003	1.036	0.964	0.898
18 months	1.004	1.023	0.974	0.917
24 months	1.004	1.015	0.982	0.934
30 months	1.003	1.009	0.988	0.948
36 months	1.002	1.006	0.992	0.959

Notes: The table reports the impulse responses of the household-specific price indices P_t^h for households at the bottom, middle, and 5% of the income distribution, and for households at the top 1% of the income distribution, expressed relative to the impulse response of the aggregate price index, P_t.

Changes in the Income Distribution and the Effectiveness of Monetary Policy

This section investigates the impact of changes in the income distribution on the effectiveness of monetary policy. With this in mind, we simulate the response of aggregate prices to a monetary shock in two counterfactual calibrations of the model with different levels of income inequality. In the first counterfactual, we reduce inequality so that the share of aggregate income held by households at the top decile of the income distribution is reduced by one half. This change in the share of income held by top households would roughly correspond to taking income inequality in the US back to 1980 levels.[9] In the second counterfactual, we increase share of income in the hands of the top decile in national income by 50 percent. In each counterfactual, we rescale the share of income of all the households below the top decile proportionally. Specifically, in the counterfactuals we set the shares to

$$s_c^h = \begin{cases} \alpha_c \times s_b^h & if\; h \in \text{top decile} \\ s_b^h \frac{1-s_c^{10\%}}{1-s_b^{10\%}} & else \end{cases},$$

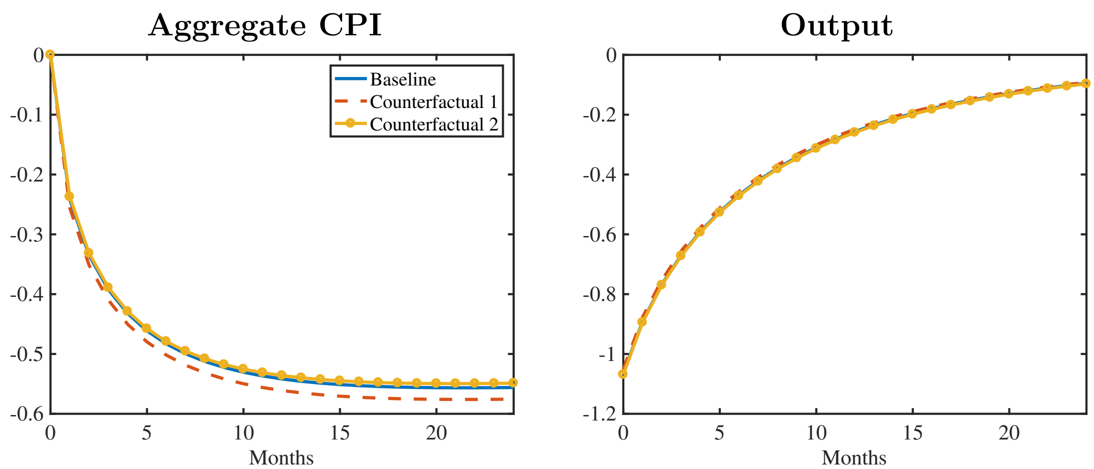
where s_b^h and s_c^h are the baseline and counterfactual shares of aggregate spending by households of type h, respectively. We set $\alpha_{c_1} = 0.5$ in the first counterfactual and $\alpha_{c_2} = 1.5$ in the second counterfactual.

To conduct these counterfactuals, we calibrate the household-specific productivities A_c^h to match the desired income distribution, while leaving aggregate income unchanged. More generally, one could imagine that the shares would also change as we change each household's income. To reflect these counterfactual changes we proceed in two steps. First, for each product category j, we split the households in the CES into consumption percentiles and perform a local polynomial regression of ω_j^h on s_b^h. Second, we calculate the predicted value of ω_j^h at the counterfactual shares s_c^h. We use these predicted values as the sectoral expenditure shares for households with counterfactual income A_c^h.

Figure 2.8 plots the impulse responses of the aggregate price index and of output in the two counterfactuals to a monetary shock that increases the nominal interest rate by 0.125 basis points on impact. The figure shows that prices are more responsive to monetary shocks in the baseline than in the counterfactual with more income inequality, and less responsive than in the counterfactual with less inequality. This is expected, since households at the top of the income distribution spend more of their income in sectors with more sticky prices. Prices decline by about 3.5% more in the counterfactual model with low income inequality for every horizon up to 24 months. However, the magnitude of the difference between the impulse responses of output is negligible. We conclude that realistic changes in inequality do not substantially alter how aggregate prices and output respond to monetary policy.

[9] Between 1980 and 2014, the share of US income held by the top 10% increased from 34% to 47%, and the share of income held by the top 1% increased from 10% to 20% according to the World Inequality Database (Alvaredo et al., 2016).

Figure 2.8 Response of the aggregate CPI and output to a monetary shock: Baseline vs. counterfactual income distributions



Notes: This figure plots the impulse responses of the aggregate price indices and output in the baseline calibration and in counterfactual a calibrations described in Section 2.5.2.

2.6 Conclusion

It has been known since at least Engel (1857, 1895) that households with different incomes consume different goods. This paper documents two novel patterns in how consumption baskets differ: in the United States, households at the top of the income distribution consume more sticky-priced goods and face substantially lower overall inflation volatility than households in the middle of the income distribution. Since the price stickiness, the volatility, and the response of prices to monetary policy differs across goods categories, these patterns suggest distributional consequences of monetary policy shocks. Because the prices of goods consumed by the high-income households are less responsive to monetary shocks, the overall CPIs of those households will react less to those shocks. We document both empirically and quantitatively that this is indeed the case. The estimated impulse responses to monetary shocks identified using the narrative approach of Romer and Romer (2004) show that CPIs of the high-income households react by about 38% less to a given monetary policy shock than CPIs of middle-income households 36 months after the shock. We then set up a multi-sector, heterogeneous-household model with sticky prices, parameterizing it to the observed sectoral heterogeneity in price stickiness and household heterogeneity in consumption baskets. In the model, the CPIs of high-income households respond 13% less to a monetary shock than the CPIs of middle-income households after 12 months.

Chapter 3

Specialization, Market Access and Real Income[1]

Co-authored with Dominick Bartelme and Andrei A. Levchenko

3.1 Introduction

The goal of this paper is to empirically estimate the effects of foreign demand and supply shocks on the real income of different countries. The notion that shocks to the size and the sectoral composition of external demand and supply lead to changes in income dates back to the origins of international economics. A voluminous theoretical literature has elaborated a number of mechanisms through which external conditions interact with the structure of domestic comparative advantage to affect real income. It has become clear that the qualitative and quantitative impacts of foreign shocks depend crucially on the strength of the various mechanisms at play, and are therefore ultimately an empirical matter.[2]

Empirical work on this question faces a number of challenges. There are many sectors and theories, but relatively few real income observations in the data. Econometric issues of endogeneity and omitted variable bias loom large. Faced with these challenges, the existing literature has coalesced around three basic approaches. One abstracts from sectoral heterogeneity altogether and focuses on the relationship between real income and the size of the external market, as determined by geography (Frankel and

[1]We thank Fernando Parro, John Ries, Andrés Rodríguez-Clare and workshop participants at several institutions for helpful suggestions. Email: dbartelm@umich.edu, tinglan@umich.edu, alev@umich.edu.

[2]Handbook chapters by Costinot and Rodríguez-Clare (2014) and Ventura (2005) review and quantify the impact of changes in openness on income levels and growth rates, respectively, under various assumptions on the structure of the economy. Rodriguez and Rodrik (2001) and Harrison and Rodríguez-Clare (2010) provide critical reviews of the empirical work on openness and income. Lederman and Maloney (2012) summarize the theoretical and empirical literature on trade patterns and income.

Romer, 1999; Redding and Venables, 2004; Feyrer, 2018). Another examines whether certain features of comparative advantage are associated with growth (e.g. Prebisch, 1959; Humphreys et al., eds, 2007; Hausmann et al., 2007). This approach abstracts from cross-country variation in external demand and supply, and lacks a common theoretical foundation. The third calibrates fully specified general equilibrium models and conducts counterfactuals (e.g. Whalley, 1985; Aguiar et al., 2016; Hsieh and Ossa, 2016). These methods deliver precise and interpretable answers, but depend heavily on the assumed model structure and a large number of parameters.

This paper develops a unified approach to quantifying the impact of foreign shocks in different sectors that strikes a balance between the clarity and rigor of structural models and more model-robust statistical methods that "let the data speak." We begin by analyzing a class of small open economy models with many sectors that satisfy four key assumptions: i) bilateral trade obeys sector-level gravity, ii) a homothetic upper-tier utility aggregator, iii) competitive goods and factor markets, and iv) a unique and smooth equilibrium mapping from the primitives to the endogenous outcomes. The production side of the economy is quite general, allowing for any number of factors, intermediate goods linkages, and external effects within and across sectors. This class contains small open economy versions of most of the quantitative trade models in the literature as special cases, including isomorphisms with various frameworks featuring monopolistic competition.

The framework delivers natural measures of sector-level foreign demand and supply shocks, which we label external firm and consumer market access respectively. These variables contain all relevant information for a country's interaction with foreign markets, and are easily estimated from the trade data using standard techniques. We employ a first order approximation to express a log change in a country's real income in terms of export and import share-weighted averages of the foreign shocks, along with domestic demand and supply shocks. The elasticities on the foreign variables measure how different foreign shocks, interacted with the domestic sectoral composition, generate different general equilibrium income impacts, thus providing a direct answer to the question posed by this paper. These elasticities also map directly to relevant parameters for trade policy.

Estimation of the model-derived equation must confront two primary challenges. The first is that there are hundreds of traded sectors and thus potentially hundreds of income elasticities that can be estimated. This is clearly not feasible given the relatively small sample of available GDP per capita data. To reduce the number of parameters to be estimated, we employ a machine learning technique to group sectors based on their characteristics into a small number of clusters. We then estimate the much smaller number of cluster-level income elasticities.

The second challenge is the common one in growth regressions: omitted variables and endogeneity. We first provide formal conditions under which the average within-cluster income elasticities are identified by an OLS regression that fully conditions on the initial equilibrium observables. The result exploits the typical invertibility properties of gravity models. To deal with the high dimensionality of the control vector we employ the Post-Double-Selection method of Belloni et al. (2014b, 2017), which is based on the approximate sparsity of the control vector to select a lower-dimensional set of "important" controls while maintaining consistency and uniformly valid inference. We rely on the fact that most countries are small in foreign markets to eliminate any direct causal from domestic shocks to foreign variables, and measure the foreign shocks in such a way as to minimize the practical relevance of this channel.

We implement our approach on UN COMTRADE trade data and decadal real income changes from the Penn World Table 9.0 over 1965-2015, with a sample of 127 countries and 268 sectors. We use the k-means clustering algorithm (MacQueen et al., 1967) along with 7 sectoral characteristics measured from US data to cluster 233 manufacturing industries into 4 clusters. It turns out that this procedure results in clusters with features that are easy to verbalize: i) processing of raw materials, ii) complex intermediate inputs, iii) capital goods, and iv) consumer goods. We group agriculture and mining sectors into their own clusters for a total of 6 clusters and therefore 12 cluster-level foreign shocks.

We find significant heterogeneity in the average impact of different foreign shocks on real income across clusters. Foreign demand shocks in complex intermediate and capital goods producing sectors have the largest impacts, with the capital goods elasticity being somewhat imprecisely estimated. Foreign demand shocks in all other sectors have small and positive income impacts. Turning to the supply shocks, we find that the largest impacts come from the capital and consumer goods sectors, although the confidence intervals are rather large. This finding reflects in part the lack of variation across countries in the identifiable foreign component of supply shocks relative to the demand shocks.

We subject our specification to robustness checks along a number of dimensions including the number of clusters, the tuning parameter used for selecting controls, measurement error in the cluster characteristics and dropping important trading partners. The most robust result is that demand shocks in complex intermediate goods have highincome elasticities and non-intermediate, non-capital goods sectors have small elasticities. The result that both supply and demand shocks in capital goods sectors have high income elasticities is moderately robust. Interestingly, when we split the sample into developed and developing countries, we find that both capital goods elasticities are much higher (and relatively precisely estimated) for developing countries across all specifications. While intriguing, the practical importance of this finding on the demand side is limited by the low shares of these goods in the export baskets of developing countries.

We conclude by examining the quantitative implications of our estimates. Given our estimated elasticities, the real income impacts are determined by the size and pattern of foreign shocks ("geography") interacted with the trade shares ("comparative advantage"). Our first exercise holds geography constant and computes the total elasticity of income with respect to uniform foreign demand and supply shocks for each country in our sample. There is substantial cross-country heterogeneity in the impacts, with rich countries benefiting more from foreign demand shocks on average due to their higher propensity to specialize in high income-elasticity sectors. Our second exercise illustrates the role of geography by holding comparative advantage constant and subjecting each country to the foreign shocks experienced by different countries in the same time period. We find that geography plays a non-trivial role in determining the growth experiences of different countries. For example, East Asian countries benefited to the tune of roughly half a percentage point of growth per year (relative to the median country) over the sample period from the rapid growth of surrounding countries, while Western European countries lost roughly half a percentage point of growth due to slow overall growth in the region.

Our paper contributes to the literature on trade and growth. A number of influential papers estimate the impact of overall openness on real income (e.g. Frankel and Romer, 1999; Rodriguez and Rodrik, 2001; Redding and Venables, 2004; Feyrer, 2018). Our paper is closer to the literature on trade patterns and income. Most of this literature studies either export or import patterns, but not both, and considers only one characteristic of trade patterns at a time. Some examples on export side include the natural resource curse literature (e.g. Humphreys et al., eds, 2007), the work on "high-income goods" (Hausmann et al., 2007), the location in the product space (Hidalgo et al., 2007), specialization in primary goods (Prebisch, 1959), or skill-intensity (Atkin, 2016; Blanchard and Olney, 2017). The literature also considered imports of capital goods (Eaton and Kortum, 2001a; Caselli and Wilson, 2004), skill-intensive goods (Nunn and Trefler, 2010; Atkin, 2016; Blanchard and Olney, 2017), or intermediate inputs (e.g. Amiti and Konings, 2007; Kasahara and Rodrigue, 2008). On the theory side, our framework is related to recent work using families of general equilibrium models to conduct trade counterfactuals (Adao et al., 2017; Allen et al., 2019; Bartelme, 2018).

The rest of the paper is organized as follows. Section 3.2 lays out the model, while Section 3.3 discusses identification and estimation. Section 3.4 describes the data and Section 3.5 presents the results. Section 3.6 discusses the quantitative implications. The details of the derivations, data construction and manipulation, and additional empirical results are collected in the Appendices.

3.2 Model

3.2.1 Economic Environment

We consider the steady state of a small open economy Home (H) in a world with N other countries (indexed by n) and K sectors indexed by k. Each sector produces a homogeneous good. Home is "small" in the sense that Home variables do not affect foreign aggregates, but it may be large in its own domestic market and will face downward sloping demand for its products in international markets (the Armington assumption).

Technology and Market Structure

There are J factors of production indexed by j, that are in fixed supply $\bar{L}_{H,j}$ and mobile across sectors. Input and output markets are competitive. Firms are infinitesimal and perceive a production technology that is constant returns to scale in their own inputs, but may feature external economies of scale that operate both within and across sectors. Given these assumptions, we can characterize the production technology in each sector by the unit cost function $c_{H,k}(\{w_{H,j}\}, \{P_{H,k}\}, \{L_{H,jk}\}, \{T_{H,k}\})$, where $\{w_{H,j}\}$ are factor prices, $\{P_{H,k}\}$ are intermediate goods prices, $\{L_{H,jk}\}$ are the factor allocations and $\{T_{H,k}\}$ are exogenous productivities. We allow this cost function to be quite general, requiring only that it is continuously differentiable. Note that we allow for cross-sectoral productivity spillovers in that the allocation of factors to other sectors may affect the unit costs in sector j.

Demand

All factor income accrues to a representative consumer. Consumers have homothetic preferences over sectoral quantity bundles $Q^C_{H,k}$.[3] Within sectors, consumers combine Home and foreign varieties in a CES fashion,

$$Q^C_{H,k} = \left(z_{H,k}^{\frac{1}{\sigma_k}} \cdot (q^C_{H,k})^{\frac{\sigma_k - 1}{\sigma_k}} + \sum_{n \in N} (q^C_{nH,k})^{\frac{\sigma_k - 1}{\sigma_k}} \right)^{\frac{\sigma_k}{\sigma_k - 1}} \tag{3.1}$$

where $z_{H,k}$ is an exogenous demand shifter. This formulation allows consumers to have home bias in consumption, so that Home products can potentially have large market share in the Home market. We assume that producers use the same aggregator for intermediate goods. We denote the sectoral CES price indices by $P_{H,k}$ and the aggregate price index by $\mathbb{P}_H$.

These assumptions on the lower tier demand functions imply a sector-level gravity equation for expenditure shares on goods from various sources. Foreign prices have two components: the source-specific costs and an iceberg bilateral component $\tau_{nH,k}$. With these assumptions, we can write the gravity equation as

$$p_{nH,k} \cdot q_{nH,k} = \frac{(c_{n,k} \cdot \tau_{nH,k})^{1-\sigma_k}}{P^{1-\sigma}_{H,k}} \cdot E_{H,k}, \quad p_{HH,k} \cdot q_{HH,k} = z_{H,k} \frac{c^{1-\sigma_k}_{H,k}}{P^{1-\sigma}_{H,k}} \cdot E_{H,k} \tag{3.2}$$

where $P^{1-\sigma}_{H,k} = z_{H,k} c^{1-\sigma_k}_{H,k} + \sum_{n \in N} (c_{n,k} \tau_{nH,k})^{1-\sigma_k}$ and $E_{H,k}$ is Home sectoral expenditure on both consumption and intermediate goods. Foreign demand for Home's commodities also takes the gravity form, with foreign imports facing some iceberg bilateral trade barriers $\tau_{Hn,k}$,

$$p_{Hn,k} \cdot q_{Hn,k} = (c_{H,k} \cdot \tau_{Hn,k})^{1-\sigma_k} \cdot \frac{E_{n,k}}{P^{1-\sigma_k}_{n,k}}. \tag{3.3}$$

We now define two key quantities. By summing export revenues across foreign export destinations, we get total foreign revenues as a function of Home costs and external *Firm Market Access* (FMA),[4]

$$\sum_{n \in N} p_{Hn,k} q_{Hn,k} = c^{1-\sigma_k}_{H,k} \cdot \underbrace{\sum_{n \in N} \tau^{1-\sigma_k}_{Hn,k} \cdot \frac{E_{n,k}}{P^{1-\sigma_k}_{n,k}}}_{FMA_{H,k}} \tag{3.4}$$

[3] We assume homotheticity in order to equate welfare with real income via a well-defined aggregate price index, which in turn allows us to make contact with national accounts data in the empirical section.

[4] This concept differs from the usual definition of market access in that it excludes the contribution of domestic demand.

Likewise, summing import expenditures across foreign sources, we get total imports as a function of Home expenditures, prices and external *Consumer Market Access* (CMA),

$$\sum_{n\in N} p_{nH,k}\cdot q_{nH,k} = \frac{E_{H,k}}{P_{H,k}^{1-\sigma_k}}\cdot \underbrace{\sum_{n\in N}(c_{n,k}\cdot \tau_{nH,k})^{1-\sigma_k}}_{CMA_{H,k}} \tag{3.5}$$

From Home's perspective, external firm and consumer market access are exogenous. Moreover, they are sufficient statistics for Home's interaction with foreign markets. Any change in foreign variables affects the Home equilibrium only through their effects on FMA and CMA.

Competitive Equilibrium

We define a competitive equilibrium in the usual way, as a set of goods and factor prices and allocations such that firms and consumers optimize taking prices as given, factor and output markets clear and trade balances. Under the assumptions above, we can characterize the equilibrium set as the set of solutions to a system of simultaneous equations in the unit cost and expenditure functions, factor prices and allocations, and trade balance (all derivations are in Appendix C.1). If factor allocations are uniquely determined given factor prices, we can further reduce the system to a set of J simultaneous equations in factor prices, equating factor supply with factor demand. Regardless of uniqueness, the set of equilibria is completely determined by the functions $c_{H,k}$ and $U(Q^C_{H,k})$, the elasticities σ_k and the exogenous variables.

Our first order approach to estimation and counterfactual welfare analysis requires a unique and smooth mapping from the exogenous variables to equilibrium outcomes. Without uniqueness the data would contain little or no information on how different foreign shocks systematically affect real income.[5] In general, without further restrictions on $c_{H,k}$ and $U(Q^C_{H,k})$ there may be multiple equilibria, with the presence of external economies being the primary culprit. It is difficult to provide sufficient conditions for uniqueness in settings with general production technology and preferences such as ours, and hence we do not pursue a characterization of the equilibrium properties of this class of models.[6] Instead, we simply assume a unique

[5]Our framework does allow small differences in either domestic fundamentals or foreign market access to have large impacts on long run real income, a feature that many models with multiple equilibria are designed to capture.

[6]Propositions 3 and 6 in Kucheryavyy et al. (2018) together provide necessary and sufficient conditions for uniqueness of an equilibrium in a labor-only small open economy with constant elasticity

and smooth equilibrium function in the relevant parameter space for the rest of the paper.

3.2.2 First Order Welfare Approximation

We now drop the H subscript to economize on notation. Our assumption of homothetic preferences equates real expenditure with welfare, while our assumption of trade balance equates nominal GDP with nominal expenditure. Thus we can write Home's welfare as

$$\frac{Y}{\mathbb{P}} = \alpha \cdot \frac{\sum_{k \in K} c_k^{1-\sigma_k} \cdot \left(z_k \frac{E_k}{P_k^{1-\sigma_k}} + FMA_k \right)}{\mathbb{P}} \tag{3.6}$$

where Y is nominal GDP and α is the share of value added in gross output. The term in the numerator of the RHS is total sales, domestic and foreign. External consumer market access enters into this expression implicitly through the sectoral price indices $P_k \equiv (z_{H,k} c_{H,k}^{1-\sigma_k} + CMA_k)^{\frac{1}{1-\sigma_k}}$.

External shocks will have two types of effects on Home's welfare in a competitive equilibrium. There will be direct effects through increased foreign sales (when FMA_k increases) and lower prices (when CMA_k increases). There will also be indirect effects as domestic producers and factor owners alter their prices and production plans and consumers alter their consumption patterns in response to these external shocks.

Our interest is in capturing the total effects of foreign shocks, both direct and indirect, in an empirical setting. To do so we make use of our assumption of a unique and smooth mapping from the domestic and foreign shocks to equilibrium quantities. Taking natural logs of Equation (3.6) and applying Taylor's theorem, the log change in real income with respect to a set of log changes in foreign shocks is approximately

$$d \ln y \approx \sum_k \delta_k^{ex} \cdot [\lambda_k^{ex} d \ln FMA_k] + \sum_k \delta_k^{im} \cdot \left[\lambda_k^{im} d \ln CMA_k\right], \tag{3.7}$$

where $y \equiv Y/\mathbb{P}$ denotes real income or welfare, λ_k^{ex} is the share of total sales accounted for by exports in sector k and λ_k^{im} is the share of total expenditures accounted for by imports in sector k.

The elasticities δ_k^{ex} and δ_k^{im} measure the total impact, direct and indirect, of foreign shocks in different industries on real income. To interpret these elasticities, consider the following natural experiment. Two small open economies, initially identical in

external scale effects.

every respect, experience a different pattern of foreign shocks. Specifically, suppose economy A experiences a 1% increase in foreign demand in industry 1 while economy B experiences a 1% increase in foreign demand in industry 2. Which economy will experience a greater real income change? Assuming both industries have the same initial export sales shares, the answer will be whichever economy gets the shock to the industry with the highest δ^{ex}. By focusing on external demand and supply shocks, rather than realized trade as in much of the literature, we can in principle separate the causal impact of external factors from that of domestic productivity or demand shifters.[7]

These elasticities are not generally "structural" parameters, except in special cases. The determinants of δ_k^{ex} and δ_k^{im} are complex and difficult to characterize analytically, because the Envelope Theorem does not apply to the competitive equilibrium even in the absence of domestic externalities, due to the economy's unexploited international market power.[8] Foreign shocks in different sectors generate different terms of trade effects, which in turn trigger different patterns of reallocation across sectors. These initial reallocations in turn generate factor price and productivity movements that imply further rounds of reallocation. These effects are especially complicated when sectors are linked through input-output relationships or productivity spillovers. Below we offer several simple examples to give some intuition for how the underlying structure of the economy determines the elasticities in different scenarios.

This very complexity provides one of the primary motivations for our approach. Rather than explicitly modeling and quantifying each aspect of the underlying structure of the economy, we aim to empirically recover the reduced form elasticities that are directly relevant to the relationship between trade and income. Our estimates will thus be robust to model uncertainty within the wide class of trade models encompassed by our framework, which offers a clear advantage over methods that require a complete specification of the model. On the other hand, we provide enough structure to enable clear interpretation, provide precise conditions for identification, and conduct local counterfactuals. These elements are missing in the reduced form literature.

There are also some costs to achieving this robustness to model uncertainty. First, fully specifying a (correct) model permits more efficient estimation of the relevant

[7]These elasticities are linked to the elasticities of real income with respect to iceberg trade costs through the identities

$$\delta_k^{ex} = \frac{1}{1-\sigma_k} \cdot \frac{\partial \ln y}{\partial \ln \tau_k^{ex}}, \quad \delta_k^{im} = \frac{1}{1-\sigma_k} \cdot \frac{\partial \ln y}{\partial \ln \tau_k^{im}}.$$

[8]See Bartelme et al. (2019) for an analysis of trade and industrial policy in an Armington SOE.

parameters. Second, a fully structural model reveals the economic mechanisms that generate the results more clearly. Third, a structural model can be solved in its non-linear form, which enables more accurate counterfactuals with respect to large shocks. We thus view our strategy as complementary to fully structural approaches.

3.2.3 Examples

Efficient Economy

We assume that the planner directly chooses quantities and factor allocations to maximize welfare, taking the production technology, factor supplies and the trade balance constraint as given. In Appendix C.1 we show that an application of the Envelope Theorem gives

$$\delta_k^{ex} = \frac{1}{\sigma_k}, \quad \delta_k^{im} = \frac{1}{\sigma_k - 1}, \quad \forall k \in K. \tag{3.8}$$

An intuition for the export elasticity comes from the fact that the optimal export tax on industry k is $1/\sigma_k$. This implies that the country earns high margins on exports from industry k, relative to another industry with the same export sales but higher σ_k. Given equal initial sales, the planner prefers a proportional increase in sales in the high margin (low σ_k) industry. The intuition for the import elasticity is a bit different: the factor $\frac{1}{1-\sigma_k}$ simply translates the increase in market access into a decrease in prices.

Single Factor Economy with No Spillovers

We now specialize our general setting to the competitive equilibrium of a single factor economy with no intermediate goods and no external economies of scale. We also assume that the upper tier utility is Cobb-Douglas. Given these assumptions,

$$\delta_k^{ex} = \kappa, \quad \delta_k^{im} = \left(\frac{1}{\sigma_k - 1} - \kappa\theta_k^d\right), \quad \forall k \in K, \tag{3.9}$$

$$\kappa = \frac{\sum_{k \in K} \lambda_k^{im}}{1 - \sum_{k' \in K}\left[\lambda_{k'}^d + (1-\sigma_{k'})\left(\lambda_{k'}^d(1-\theta_{k'}^d) + \lambda_{k'}^{ex}\right)\right]}.$$

where λ_k^d is the initial share of domestic sales in industry k in total sales, and $\theta_k^d = \frac{\lambda_k^d}{\lambda_k^d + \lambda_k^{im}}$. Unlike the case of an efficient economy, here the export elasticity is constant across industries. This is because in a single factor economy without spillovers, labor allocations to exports in each industry are proportional to the export sales share λ_k^{ex}. This implies that the indirect effect (through the wage) of a shock to $\ln FMA_k$

is proportional to the export share; since the direct effect is also proportional to the export share the overall effect is proportional as well. The constant of proportionality κ reflects the overall importance of trade to the economy as well as the distribution of sales across foreign and domestic customers and their covariance with the trade elasticities. The import elasticity is modified (relative to the efficient case) to account for the negative impact of foreign competition on domestic producers.

Single Factor Economy with Industry Spillovers

We now augment the single factor economy above with endogenous within-industry productivity spillovers as in Kucheryavyy et al. (2018) and Bartelme et al. (2019), so that $c_k = \frac{w}{T_k L_k^{\gamma_k}}$. To simplify the analysis we assume a Cobb-Douglas upper tier and zero domestic sales, as well as the condition $\gamma_k(\sigma_k - 1) < 1, \;\; \forall k$ to ensure a unique interior equilibrium. The elasticities are now given by

$$\delta_k^{ex} = \kappa \cdot \frac{1}{1 - \gamma_k(\sigma_k - 1)}, \quad \delta_k^{im} = \left(\frac{1}{\sigma_k - 1}\right), \quad \forall k \in K, \qquad (3.10)$$
$$\kappa = \frac{1}{1 - \sum_{k' \in K} \frac{(1+\gamma_{k'})(1-\sigma_{k'})}{1-\gamma_{k'}(\sigma_{k'}-1)} \lambda_{k'}^{ex}}.$$

All else equal, foreign demand shocks in sectors with larger productivity spillovers generate a higher income change. Notice that, for a given γ_k, higher σ_k also implies a higher income elasticity. This reflects the fact that scale economies are more valuable in sectors with more elastic international demand; in less elastic sectors, achieving higher productivity comes at the expense of significantly lower export prices.[9]

3.2.4 Isomorphisms and Extensions

We have derived our results using the competitive equilibrium of an Armington economy to maximize clarity and simplicity. However, the crucial assumptions are the gravity assumption on trade flows, homothetic upper tier preferences and the unique equilibrium mapping that validates our first order approach. Thus models with alternative micro-foundations for gravity, such as those based on Eaton and Kortum (2002), Krugman (1980), or Melitz (2003) with a Pareto distribution for productivity, will be isomorphic to our model in the sense that they have a first order approximation

[9] Our assumption of zero domestic sales implies that foreign supply shocks do not affect domestic prices or production decisions. With positive domestic sales the formulas become quite messy, but the general intuition is still that countries prefer demand shocks in high $\gamma_k(\sigma_k - 1)$ sectors and prefer supply shocks in low $\gamma_k(\sigma_k - 1)$ sectors.

of the same form as Equation (3.7) and the same interpretation of the market access elasticities.

Our framework is static, and thus should be interpreted as capturing long run differences across steady states. Our assumption of fixed factor endowments formally rules out dynamic models of factor accumulation, but we can extend our approach to allow for this feature as well by letting the steady state factor supplies depend on the other exogenous variables of the model through long run factor supply equations.

3.3 Identification and Estimation

3.3.1 Identification

We now consider identification of the elasticities δ_k^{ex} and δ_k^{im} based on Equation (3.7). To match our empirical setting, we consider a world populated by many small open economies (indexed by i) over many time periods (indexed by t), with a fixed set of industries indexed by k. The log change in real income in country i between time t and $t+1$ is approximately

$$d\ln y_{i,t} \approx \sum_k \delta_{ik,t}^{ex} \cdot \left[\lambda_{ik,t}^{ex} d\ln FMA_{ik,t}\right] + \sum_k \delta_{ik,t}^{im} \cdot \left[\lambda_{ik,t}^{im} d\ln CMA_{ik,t}\right] + \sum_k \delta_{ik,t}^{T} \cdot d\ln T_{ik,t} \tag{3.11}$$

where $d\ln x_{ik,t} = \ln x_{ik,t+1} - \ln x_{ik,t}$ for $x = FMA, CMA, T$.[10]

The variables $d\ln FMA_{ik,t}$, $d\ln CMA_{ik,t}$ and $d\ln T_{ik,t}$ on the right hand side of this equation are not directly observable. However, $FMA_{ik,t}$ and $CMA_{ik,t}$ can be consistently estimated using conventional gravity equation techniques (Head and Mayer, 2014). We defer a detailed discussion of our estimation strategy for these variables to Section 3.4, and assume that they are known with certainty for the remainder of this section. In contrast, the domestic productivity shocks $T_{ik,t}$ cannot be observed or estimated without knowledge of the full model structure. We treat the domestic shocks as unobservable, which leads to the empirical specification

$$d\ln y_{i,t} = \nu_t + \sum_k \delta_{ik,t}^{ex} \cdot \left[\lambda_{ik,t}^{ex} d\ln FMA_{ik,t}\right] + \sum_k \delta_{ik,t}^{im} \cdot \left[\lambda_{ik,t}^{im} d\ln CMA_{ik,t}\right] + \epsilon_{i,t}, \tag{3.12}$$

[10]For expositional purposes we assume that neither the factor supplies nor the domestic demand shifters change. It is straightforward but notationally cumbersome to add these terms. All our results regarding identification in the presence of unobserved productivity shocks apply to these variables as well.

where ν_t is the mean time-t domestic shock term, and $\epsilon_{i,t} = \sum_k \delta^T_{ik,t} \cdot d \ln T_{ik,t} - \nu_t$.

As written, equation (3.12) has a larger number of parameters ($2K \times N \times T$) than observations ($N \times T$). In some simple examples the elasticities depend only on industry characteristics, but in general they also depend on the initial equilibrium (the point of approximation) and are thus country and time-specific as well. This issue is compounded by the fact that we observe a large number of distinct traded industries relative to the number of medium-run country-time growth rates in the sample, making even the estimation of industry-specific elasticities problematic in our finite sample.

We begin by clustering "similar" industries together, where similarity is defined as closeness in the space of industry characteristics. We measure a number of industry characteristics that are likely to affect the elasticities, then cluster the industries using the k-means algorithm commonly used in machine learning and statistics. Section 3.4 describes the industry characteristics and the clustering algorithm in detail. For now, simply suppose that we have arrived at some clustering scheme $g \in G$. Using this notation, we can rewrite Equation (3.12) as

$$d \ln y_{i,t} = \nu_t + \sum_{g \in G} \delta^{ex}_g \cdot [d \ln FMA_{ig,t}] + \sum_{g \in G} \delta^{im}_g \cdot [d \ln CMA_{ig,t}] + \mu_{i,t} + \epsilon_{i,t}, \quad (3.13)$$

where

$$\delta^{ex}_g = \frac{1}{K_g} \sum_{k \in g} E_{i,t}[\delta^{ex}_{ik,t}], \quad d \ln FMA_{ig,t} = \sum_{k \in G} \lambda^{ex}_{ik,t} d \ln FMA_{ik,t}, \quad (3.14)$$

$$\mu_{i,t} = \sum_{g \in G} \sum_{k \in g} (\delta^{ex}_{ik,t} - \delta^{ex}_g) \left[\lambda^{ex}_{ik,t} d \ln FMA_{ik,t}\right] + \sum_{g \in G} \sum_{k \in g} (\delta^{im}_{ik,t} - \delta^{im}_g) \left[\lambda^{im}_{ik,t} d \ln CMA_{ik,t}\right], \quad (3.15)$$

K_g is the number of industries in g and similar definitions apply to δ^{im}_g and $d \ln CMA_{ig,t}$. The parameters of interest are the δ_g's, which are the within-cluster average of the average partial effects $E_{i,t}[\delta_{ik,t}]$. They can be interpreted as the best guess for the real income impact of a unit shock to log market access in industry $k \in q$ for a randomly chosen country and time period, conditional only on the identity of the cluster.

Identification requires the conditional independence of the foreign shocks and the two error components, $\mu_{i,t}$ and $\epsilon_{i,t}$. As it stands, Equation (3.13) does not satisfy this condition, since both the foreign shocks and the error components depend on the initial equilibrium. The foreign shocks are obviously functions of the initial equilibrium, via the trade share weights $\lambda^{ex}_{ik,t}$ and $\lambda^{im}_{ik,t}$. Less obviously, the error components $\mu_{i,t}$ and $\epsilon_{i,t}$ are also functions of the initial equilibrium. This dependence stems from several sources, primarily the dependence of the country-industry-time-specific elasticities $\delta_{ik,t}$ on the initial equilibrium and any serial correlation in the domestic shocks $d\ln T_{ik,t}$. Intuitively, the identification challenge is to ensure that "all else is equal" across countries receiving different "treatments," i.e. different patterns of foreign shocks. Note that the large number of potential channels for correlation between the errors and the independent variables makes it impossible to sign the bias that would arise from estimating Equation (3.13) using OLS.

This discussion suggests that we could identify the cluster-level average treatment effects if we condition on all relevant information on the initial equilibrium. We exploit the structure of the gravity to rigorously show how we can do so. Recall from Section 3.2 that we assume the existence of a smooth and one-to-one equilibrium map which determines every endogenous variable, including the $\delta_{ik,t}$, as a function of the set of exogenous variables $\{\{T_{ik,t}\},\{z_{ik,t}\},\{FMA_{ik,t}\},\{CMA_{ik,t}\},\{\bar{L}_{ij,t}\}\}$. In principle, the $FMA_{ik,t}$, $CMA_{ik,t}$ and $\bar{L}_{ij,t}$ are all observable while the domestic supply and demand shifters $T_{ik,t}$ and $z_{ik,t}$ are not. However, gravity models of trade typically have the property that, conditional on the rest of the exogenous variables and the parameters of the model, the trade flows $\lambda^{ex}_{ik,t}\cdot Y_{i,t}$ and $\lambda^{im}_{ik,t}\cdot E_{i,t}$ can be inverted to recover the $T_{ik,t}$ and $z_{ik,t}$ that generated them. We assume that the underlying model has this property as well, which allows us to characterize any variable in the initial equilibrium as functions of observables. Once we condition on the initial equilibrium via these observables, identification of the elasticities follows, provided that the residual innovations in domestic productivity and demand are uncorrelated with the foreign shocks. Our small open economy assumption makes this identification condition internally consistent with our model in the sense that there can be no direct causal relationship between the domestic and foreign shocks, and thus it involves only restrictions on the joint distribution of the exogenous variables.[11]

We now provide formal sufficient conditions for identification for two special cases of the general model in Equation (3.13), then discuss the general case. Our discussion assumes that the mapping from the initial equilibrium observables to the unobservables is sufficiently smooth to be well approximated by linear combinations of functions of initial observables, such as dummies, polynomials, splines, and interactions. We denote the (potentially high dimensional) vector of approximating variables by $\mathbf{w}_{i,t}$, and WLOG assume that each component has mean zero.

[11]In a large economy, domestic shocks will affect foreign variables. We measure our foreign shocks so as to minimize the effect of any violations of this assumption in the data, and conduct robustness checks with respect to this assumption in Section 3.5.

Constant Treatment Effects Within Clusters

In this case, the elasticities are constant within cluster, i.e. $\delta^{ex}_{ik,t} = \delta^{ex}_{g}$ and $\delta^{im}_{ik,t} = \delta^{im}_{g}$. Under this assumption, we can write Equation (3.13) as

$$d \ln y_{i,t} = \nu_t + \sum_{g \in G} \delta^{ex}_g \cdot [d \ln FMA_{ig,t}] + \sum_{g \in G} \delta^{im}_g \cdot [d \ln CMA_{ig,t}] + \boldsymbol{\eta} \mathbf{w}_{i,t} + \tilde{\epsilon}_{i,t}, \quad (3.16)$$

$$\tilde{\epsilon}_{i,t} = \sum_{k \in K} \boldsymbol{\eta}_k \mathbf{w}_{i,t} \cdot \xi^T_{ik,t}, \quad E[\tilde{\epsilon}_{i,t}] = 0, \quad E[\xi^T_{ik,t} | \mathbf{w}_{i,t}] = 0 \ \ \forall k.$$

Here the $\xi^T_{ik,t}$ are the component of the $d \ln T_{ik,t}$ that is unforecastable by the initial equilibrium variables $\mathbf{w}_{i,t}$. Then a sufficient condition for an OLS regression that controls for $\mathbf{w}_{i,t}$ to identify the δ_g's is that the conditional expectation of the productivity innovations with respect to the foreign shocks and controls is zero,

$$E_{i,t}[\xi^T_{ik,t} | \mathbf{w}_{i,t}, \{d \ln FMA_{ig,t}\}, \{d \ln CMA_{ig,t}\}] = 0, \quad \forall k. \quad (3.17)$$

This condition implies that once we control for the initial equilibrium, the foreign shocks vary independently from the domestic shocks and thus provide exogenous variation that can be leveraged for identification.

Constant Treatment Effects Within Cluster-Country-Time

Our identification result above assumed away the problem of inference in the presence of heterogeneous treatment effects within clusters. We now allow the treatment effects to vary by country and time period, but not across sectors within a given cluster-country-time, i.e. $\delta^{ex}_{ik,t} = \delta^{ex}_{ig,t}$ and $\delta^{im}_{ik,t} = \delta^{im}_{ig,t}$. Unlike the typical application, the heterogeneity in our treatment effects is not random after conditioning on the initial equilibrium. However, we can fully control for the remaining dependence using interactions of the initial equilibrium variables with the treatments. Formally, let $\mathbf{s}_{i,t}$ denote the vector of interactions between the initial equilibrium variables $\mathbf{w}_{i,t}$ and the k-level foreign shocks $d \ln FMA_{ik,t}$ and $d \ln CMA_{ik,t}$. Then we can write Equation (3.13) as

$$d \ln y_{i,t} = \nu_t + \sum_{g \in G} \delta^{ex}_g \cdot [d \ln FMA_{ig,t}] + \sum_{g \in G} \delta^{im}_g \cdot [d \ln CMA_{ig,t}] + \boldsymbol{\eta} \mathbf{w}_{i,t} + \boldsymbol{\phi} \mathbf{s}_{i,t} + \tilde{\epsilon}_{i,t}, \quad (3.18)$$

$$\tilde{\epsilon}_{i,t} = \sum_{k \in K} \boldsymbol{\eta}_k \mathbf{w}_{i,t} \cdot \xi^T_{ik,t}, \quad E[\tilde{\epsilon}_{i,t}] = 0, \quad E[\xi^T_{ik,t} | \mathbf{w}_{i,t}] = 0 \ \ \forall k.$$

Once we control for both the initial equilibrium and the dependence of the individual treatment effects on the initial equilibrium, our condition for identification remains the same as in the constant elasticity case. Note that our de-meaning of $\mathbf{w}_{i,t}$ ensures that there is not full collinearity between the cluster-level treatments and the control $\mathbf{s}_{i,t}$.

General Treatment Effects

We now examine the case where the treatment effects also vary by industry within each country-time-cluster. Here we face a more difficult challenge to identification: the mean treatment effects by industry within a cluster vary in a way that we cannot control for without introducing collinearity with the treatments. Formally, and with a slight abuse of notation, let $\mathbf{s}_{i,t}$ now denote the vector of interactions between the initial equilibrium variables $\mathbf{w}_{i,t}$ and the k-level foreign shocks $\lambda^{ex}_{ik,t} d\ln FMA_{ik,t}$ and $\lambda^{im}_{ik,t} d\ln CMA_{ik,t}$. Then we can write Equation (3.13) as

$$\begin{aligned} d\ln y_{i,t} =& \nu_t + \sum_{g\in G} \delta^{ex}_g \cdot [d\ln FMA_{ig,t}] + \sum_{g\in G} \delta^{im}_g \cdot [d\ln CMA_{ig,t}] + \boldsymbol{\eta}\mathbf{w}_{i,t} + \boldsymbol{\phi}\mathbf{s}_{i,t} + \tilde{\epsilon}_{i,t}, \quad (3.19) \\ \tilde{\epsilon}_{i,t} =& \sum_{k\in K} \boldsymbol{\eta}_k \mathbf{w}_{i,t} \cdot \xi^T_{ik,t} + \sum_{g\in G}\sum_{k\in g} (\delta^{ex}_k - \delta^{ex}_g)\lambda^{ex}_{ik,t} d\ln FMA_{ik,t} \\ &+ \sum_{g\in G}\sum_{k\in g} (\delta^{ex}_k - \delta^{ex}_g)\lambda^{im}_{ik,t} d\ln CMA_{ik,t}, \\ E[\tilde{\epsilon}_{i,t}] =& 0, \quad E[\xi^T_{ik,t}|\mathbf{w}_{i,t}] = 0 \ \ \forall k, \end{aligned}$$

where δ^{ex}_k and δ^{im}_k are the mean treatment effects at the industry level. Since the q-level treatments are just sums of the k-level treatments, there is a structural correlation between the error term and the treatments that may lead to bias.

Intuitively, the source of the bias comes from the potential for certain sectors to contribute disproportionately to the variation of the cluster level treatment, either because they comprise a larger share of trade or because they face more volatile foreign shocks. If that is the case, then the estimated cluster-level mean treatment effects will disproportionately reflect the contributions of those more highly weighted sectors. As an extreme example, suppose that in a given cluster with 100 industries, only one industry ever experiences a foreign shock. Clearly we cannot use any amount of data to recover the cluster-level mean treatment effect; what we will recover instead is the mean treatment effect for that industry.[12] In the more general case, the elasticities

[12]To further build intuition, it may be helpful to consider the following special case in which there

that we recover will be weighted averages of the industry-level mean treatment effects, where the weights reflect the likelihood of treatment conditional on the controls.

3.3.2 Estimation

We have shown that the group-level treatment effects are identified under reasonable conditions once we adequately control for the initial equilibrium observables. However, the vector of controls may be quite high-dimensional relative to the sample size. This is certainly the case in our application, where we have hundreds of medium-term growth rates but thousands of controls if we include initial import and export shares, interactions, etc. Thus conventional OLS estimation is infeasible.

To address this issue, we use the Post-Double-Selection estimator developed by Belloni et al. (2014b, 2017). This approach involves selecting a subset of "important" controls by regressing each dependent and independent variable on the full set of potential controls using an estimator that sets some or all of the coefficients to zero (e.g. LASSO). The selection is "double" in that the controls are selected based on their correlations with both the dependent and independent variables. The union of the sets of controls that are thus selected (i.e. have non-zero coefficients) in each regression then form the control set for an OLS regression of the dependent variable on the independent variables.

Belloni et al. (2014b) show that this estimator is consistent and asymptotically normal, with the usual standard errors generating uniformly valid confidence intervals, under conditions that are quite plausible in our setting. The most important condition is that the true control vector admits an *approximately sparse* representation in the sense that the true control function can be well-approximated by a function of a subset of the controls.[13] This condition does not require that the control function exhibit true sparsity, only some combination of true sparsity, many small coefficients, and high correlation between controls. These conditions seem reasonable in our setting.

is no bias: trade shares are constant within clusters for any given country-time period, and the changes in foreign market access are *i.i.d.* within cluster-country-time period.

[13]We refer the reader to Belloni et al. (2014a), Belloni et al. (2014b) and Belloni et al. (2017) for additional details and regularity conditions.

3.4 Data, Clustering and Foreign Shock Estimation

This section briefly summarizes our data sources and measurement strategy. Appendix C.2 collects the detailed descriptions of all steps.

3.4.1 Data

Our empirical implementation requires data on (i) real income per capita, (ii) sectoral bilateral trade flows and trade barriers, and (iii) sectoral characteristics. Income per capita is sourced from the Penn World Tables 9.0, computed as the real GDP at constant national prices divided by population. We drop countries with population less than 2 million from our sample. Per capita income growth is computed at 10-year intervals for a maximum of 5 ten-year growth rates per country (there are some missing values).

The bilateral trade flow data at the 4-digit SITC Rev 2 level come from the UN Comrade Database. We convert the trade data from the SITC to the 1997 NAICS classification. Appendix C.2.1 describes the construction of the concordance in detail. All in all, the 784 4-digit SITC items are matched to 268 NAICS sectors. Among them are 233 manufacturing, 26 agricultural, and 9 mining sectors. Geographic variables (bilateral distance and contiguity measures) come from CEPII. The final sample covers 127 countries, 268 sectors and 5 decades from 1965 to 2015, with a total of 548 10-year GDP growth rate observations.

The 233 manufacturing sectors are grouped into clusters based on their sectoral characteristics. We use data from the United States to measure the sectoral characteristics, since sectoral data at a comparable 4-digit level of sectoral disaggregation are not available for a large sample of countries. We collect data on 7 sectoral features: investment sales shares, intermediates using shares, intermediates sales shares, 4-firm concentration ratios, skilled worker shares, physical capital intensities, and the contract intensity of inputs. Sectoral characteristic variables are collected from various data sources with similar but not always identical industry classifications. We convert all of them to the 1997 NAICS classification.

Our measures of the investment sales shares, intermediates sales shares and intermediate using shares are based on data from the 1997 Benchmark Detailed Make and Use Tables. The investment sales share is computed as the ratio of spending on sector k for investment purposes to the the total gross output of sector k. Thus, this variable captures in a continuous way the extent to which sector k produces capital goods. Similarly, intermediates sales and using shares of gross output capture the extent to which sector k is a large producer or user of intermediate goods, respectively. The four-firm concentration ratios are sourced from the 2002 Economic Census. The skilled worker shares are calculated as the share of workers in sector k that have a bachelor degree or higher, and are computed based on data from the 2000 American Community Survey. The capital intensity variable is measured as 1 minus the labor share of value added (payroll), based on the NBER-CES Manufacturing Industry Database. The contract intensity of a sector is measured as the fraction of a sector's inputs that need relationship-specific investments, and comes from Nunn (2007). We use the version of this variable that measures the fraction of inputs not sold on organized exchanges and not reference priced to capture the importance of relationship-specific investments in a sector.

3.4.2 K-means Clustering

As discussed above in Section 3.3, given our sample size and the large number of industries, we focus on estimating average treatment effects within groups or clusters of industries. While average treatment effects for any set of industry groups are identified, it is more useful and interesting to group industries according to characteristics that are both observable and related to the treatment effects. We implement this approach by measuring the 7 characteristics (described in the previous subsection) for each industry, then assigning industries to clusters based on their proximity in the space of characteristics. We apply this approach to the manufacturing industries in our sample.

We use the k-means clustering algorithm (MacQueen et al., 1967) to group sectors into clusters. Sectors are assigned to clusters based on their characteristics so as to minimize the within-cluster sum of squared deviations from the cluster mean. The k-means algorithm works as follows: given m manufacturing sectors, each with a vector of n different sectoral characteristics, $x^{(i)} \in \mathbb{R}^n$, $i = 1, \ldots, m$, assign the m sectors into G clusters. The G clusters are labeled as $g = 1, 2, \ldots, G$.

i. Initialize cluster centroids $\mu_1, \mu_2, \ldots, \mu_G$ for each cluster.

ii. Assign each sector $x^{(i)}$ to closest cluster centroids. The cluster assignment is

$c(i) \in \{1, 2, \ldots, G\}$,

$$c(i) = \underset{g \in \{1,\ldots,G\}}{\operatorname{argmin}} ||x^{(i)} - \mu_g||^2.$$

iii. Replace cluster centroid μ_g by the coordinate-wise average of all points (sectors) in the gth cluster,

$$\hat{\mu}_g = \frac{\sum_{i=1}^{m} \mathbf{1}(c(i) = g) \cdot x^{(i)}}{\sum_{i=1}^{m} \mathbf{1}(c(i) = g)}.$$

iv. Iterate on steps 2 and 3 until convergence.

We use the "k-means ++" algorithm proposed by Arthur and Vassilvitskii (2007) to choose the initial values for the k-means clustering algorithm, and do extensive checks using alternative starting points. Following standard practice, we normalize the values of each characteristic to have zero mean and unit variance.[14]

The algorithm above requires a choice of the number of clusters. There is no unambiguously optimal method, although there are a number of conceptually similar approaches based on maximizing various measures of cluster fit with respect to the number of clusters. We use the silhouette width (Rousseeuw, 1987) as our measure of cluster fit. Loosely speaking, the silhouette width measures how similar industries within a cluster are to each other relative to industries in the nearest cluster. A good clustering scheme will maximize the average silhouette width while minimizing the number of sectors near the boundaries. The silhouette analysis suggests that either 4 or 5 are good values for number of clusters. Appendix C.2.2 reports the results of the silhouette analysis along with a fuller discussion. In the interest of parsimony we choose to group the 233 manufacturing industries into 4 clusters in our baseline analysis, and show that our results are insensitive to this choice in Appendix C.2.2.

Table 3.1 summarizes the characteristics of the 4 clusters. Since each cluster has some salient features that distinguish it from others, we name the clusters based on these key features. It is important to stress that the clustering procedure does not produce these cluster labels, nor does our identification strategy hinge upon them. We use the cluster names (shown in the last row of Table 3.1 purely for expositional purposes. Note that there is no information contained in cluster numbers (1, 2, ...).

The sectors in cluster 1 have the highest intermediate sales and using shares, and lowest contract intensity. We label these sectors "raw materials processing" sectors.

[14]This step is prudent because k-means clustering is not invariant to the scale used to measure the characteristics. If a particular characteristic takes on a broader range of values than the others, it will be given higher weight when assigning industries to clusters.

These sectors typically involve the first stage of turning raw materials into manufactured goods. Cluster 2 has the second-highest intermediate sales shares (after cluster 1), but considerably higher contract intensity than cluster 1. We thus label it "complex intermediates." Cluster 3 stands out most clearly as capital goods, with an average investment share of 0.52 compared to investment shares ranging from 0.00 to 0.05 in the other clusters. Cluster 4 has nearly the lowest average intermediate sales share, and a negligible average investment sales share. Thus we label it "consumer goods." Table C.1 in Appendix C.2.2 lists the 3 most representative sectors in each cluster, defined as those closest to the cluster centroid.

As we do not have information on these characteristics for non-manufacturing sectors, we group all agricultural sectors to Cluster 5, and all mining sectors to Cluster 6. In total, the 268 sectors are grouped into 6 clusters.

Table 3.1 Summary Statistics of Clusters in Manufacturing

	cluster					
	1	2	3	4	Mean	Std. Dev.
Inv. Share	0.00	0.05	0.52	0.04	0.13	0.22
Int. Using	0.78	0.58	0.65	0.66	0.66	0.16
Int. Sales	0.84	0.70	0.27	0.28	0.57	0.31
Conc. Ratio	0.47	0.27	0.38	0.56	0.40	0.21
Sk. Share	0.32	0.28	0.35	0.36	0.32	0.13
Cap. Int.	0.68	0.55	0.54	0.70	0.61	0.10
Con. Int.	0.26	0.56	0.73	0.52	0.51	0.22
Num of ind.	60	84	47	42		
Trade share	0.33	0.26	0.23	0.11		
Label	Raw Materials Processing	Complex Intermediates	Capital Goods	Consumer Goods		
Abbreviation	RAW	INT	CAP	CONS		

Notes: This table reports the summary statistics of the sectoral characteristics among the sectors selected into each cluster. The last two columns report the mean and standard deviations of those characteristics among all manufacturing sectors. The row "Num. of ind" reports the number of sectors in each cluster, and "Trade share" reports the fraction of world trade accounted for by sectors in that cluster. The bottom panel lists the intuitive labels of the clusters, as well as 3-letter abbreviations. Both are heuristic and assigned by the authors.

3.4.3 Estimation Strategy for $FMA_{ik,t}$ and $CMA_{ik,t}$

To obtain $FMA_{ik,t}$ and $CMA_{ik,t}$ for country i sector k at time t, we estimate structural sector-specific gravity equations using the matrix of sectoral bilateral trade flows at decadal intervals.[15] For a given sector k at time t, the gravity equation (3.2) can be rewritten as

$$\lambda_{ink,t} = c_{ik,t}^{1-\sigma_k} \cdot P_{nk,t}^{\sigma_k - 1} \cdot \tau_{ink,t}^{1-\sigma_k}, \tag{3.20}$$

where $\lambda_{ink,t}$ denotes the share of n's expenditure on sector k that is sourced from country i. Since we do not observe domestic trade flows, we calculate $\lambda_{ink,t}$ as the share of import expenditure. We model the bilateral resistance term $\tau_{ink,t}^{1-\sigma_k}$ as a function of geographic distance and contiguity with sector-time-specific coefficients, leading to our empirical specification

$$\lambda_{ink,t} = \kappa_{ik,t}^{ex} \cdot \kappa_{nk,t}^{im} \cdot Distance_{in}^{\zeta_{kt}} \cdot \exp\left(\xi_{kt} \cdot Contig_{in}\right) \cdot \epsilon_{ik,t}, \tag{3.21}$$

where $\kappa_{ik,t}^{ex}$ is the exporter fixed effect, $\kappa_{nk,t}^{im}$ is the importer fixed effect, ζ_{kt} and ξ_{kt} are the distance and common border coefficients. We estimate the non-linear equation (3.21) using the Poisson Pseudo-Maximum Likelihood (PPML) method proposed by Silva and Tenreyro (2006) and Eaton et al. (2012), separately for every sector and time period.

We use our estimates from equation (3.21) to construct the external market access terms as follows:

$$FMA_{ik,t} = \sum_{n \neq i} E_{nk,t} \cdot \kappa_{nk,t}^{im} \cdot Distance_{in}^{\zeta_{kt}} \cdot \exp\left(\xi_{kt} \cdot Contig_{in}\right) \tag{3.22}$$

$$CMA_{nk,t} = \sum_{i \neq n} \kappa_{ik,t}^{ex} \cdot Distance_{in}^{\zeta_{kt}} \cdot \exp\left(\xi_{kt} \cdot Contig_{in}\right), \tag{3.23}$$

where $E_{nk,t}$ is n's total foreign expenditure in k at time t.

In practice, we add two wrinkles to the method described above. First, we remove any direct effect of a country's exports and imports on the fixed effects of their trading partners by estimating equation (3.21) N times for each sector and time period, each time leaving out the trade flows from a particular country i. We then construct each country i's foreign shocks using the estimates from the regression that omitted its data. Second, as is well known, $\kappa_{ik,t}^{ex}$ and $\kappa_{nk,t}^{im}$ are identified only up to a sector-time-specific multiplicative constant and require normalization. Rather than the usual practice of designating a particular numéraire country, we restrict the sum of the logged importer effects to be zero. This normalization ensures that the relative growth rates of the foreign shocks across industries are not driven by fluctuations in the trade flows of the numéraire country, minimizing measurement error. Appendix C.2.3 provides a

[15] To reduce measurement error, we use three year averages of the trade flows.

detailed discussion.

This procedure uses only foreign data to construct external market access and projects bilateral flows onto a small number of variables (distance and contiguity). By construction, it excludes domestic factors that act as country-specific average export taxes (on FMA) or average import taxes (on CMA).[16] It also excludes idiosyncratic bilateral factors that affect trade flows. This tends to minimize concerns about domestic policies or shocks influencing measured market access, but does introduce some measurement error.

3.5 Empirical Results

3.5.1 Baseline Estimates

The top panel of Figure 3.1 presents the estimation results graphically, by displaying the coefficients on the foreign demand shocks in the left panel, and for the foreign supply shocks in the right panel, by cluster. Clusters 1-4 are manufacturing clusters obtained by the k-means algorithm, cluster 5 is agriculture, and cluster 6 mining and quarrying. The bars depict 95% confidence intervals, obtained with standard errors clustered at the country level. The specification includes the log of initial GDP per capita.

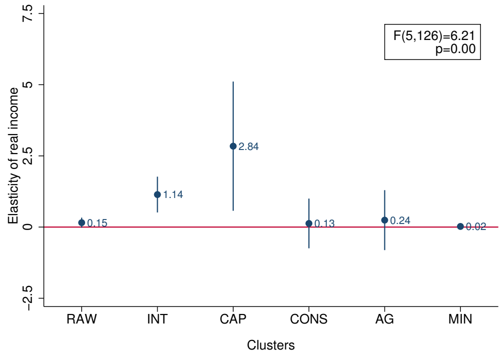
The first apparent feature of the results is the considerable heterogeneity in the coefficients. Indeed, the F-tests reject the equality of these coefficients at the 1% level of significance. When it comes to foreign demand shocks, two clusters stand out: export opportunities in cluster 2 ("complex intermediates", labeled "INT"), and cluster 3 ("capital goods", or "CAP") seem to have a larger and statistically significant positive real income effect than the other clusters, although the confidence interval on CAP is wide.

On the foreign supply shock side, there is also some heterogeneity in the coefficients (equality is rejected at the 1% level), but only the shock to the consumer goods supply exhibits a statistically significant positive impact on income. Overall, the foreign supply shocks have both much larger magnitudes and standard errors. The latter feature makes it challenging to draw sharp conclusions about the impact of foreign supply shocks on income. In practice, the variation in the FMA terms is an order of

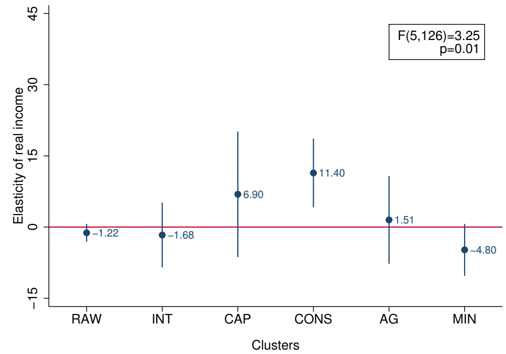
[16]For example, if a country's trading partners lowered their prices to the world by 10% but its government raised import tariffs by the equivalent amount, our procedure would record an increase in external CMA but it would not affect the prices experienced by consumers.

Figure 3.1 Cluster-Specific Coefficients and Confidence Intervals

I. OLS Estimates

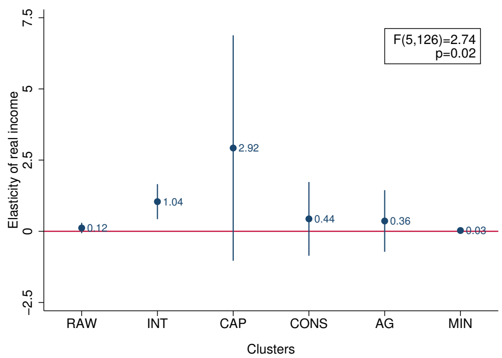


(a) Foreign Demand Shocks

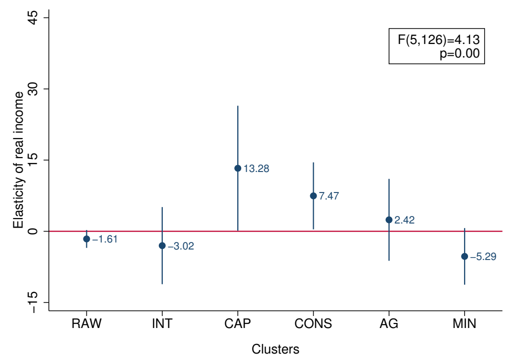


(b) Foreign Supply Shocks

II. LASSO Estimates



(c) Foreign Demand Shocks



(d) Foreign Supply Shocks

Notes: This figure reports the coefficients in estimating Equation (3.13), for the foreign demand shocks (FMA) (left panels), and foreign supply shocks (CMA) (right panels). All specifications control for initial GDP per capita. The top panel displays the baseline OLS estimates. The bottom panel displays the post double-LASSO estimates. 16 control variables are selected in the double-selection step. The bars display the 95% confidence bands, that use standard errors clustered by country. The boxes display the results of an F-test for equality of the coefficients in each plot.

magnitude larger than the variation in CMA terms. This is sensible from an economic standpoint: examination of the functional forms for FMA and CMA in equations (3.22) and (3.23) reveals that foreign demand shocks are determined by both changes in foreign prices/costs as well as changes in the overall foreign expenditure. On the other hand, foreign supply shocks are driven purely by changes in foreign costs. As a result, the FMA terms have much greater variation in the data. Statistically, it is thus not surprising that a regressor with a smaller standard deviation has a higher point estimate. The large standard errors, however, imply a relative lack of confidence in those estimates.

The bottom panel of Figure 3.1 displays the Post-Double-Selection estimation results (Belloni et al., 2014b). The procedure is described in detail in Appendix C.2.4.

The specification includes a full set of potential controls, namely the industry-level initial equilibrium variables (initial import and export shares, weighted initial firm and consumer market access levels, the squares, and the interactions), interactions between the initial equilibrium variables and the industry-level foreign shocks, initial capital, and initial real GDP per capita. In total, 3219 potential control variables are included and 14 of them are selected in the double-selection procedure via LASSO. Appendix Table C.3 lists the selected controls in the Post-Double-Selection estimation.[17] Substantively the results are quite similar to the OLS specification, although some confidence intervals widen. Foreign demand shocks in complex intermediate and capital goods retain a large estimated impact on income, while all other sectors have low estimated impacts. The results for the foreign supply shocks are also similar to the OLS, with a somewhat larger estimate for the capital goods sector and a somewhat smaller one for the consumer goods sector. Once again, however, the confidence intervals are quite wide.

3.5.2 Robustness Checks

Assignment of Sectors to Clusters

One concern with our approach is that clusters may be fragile due to some sectors being on the margins between clusters. If those sectors are particularly influential, then the results could be sensitive to the assignment of specific sectors to clusters. To assess the role of marginal sectors in our results, we perform two exercises. First, we add a 5th manufacturing cluster. The results of re-clustering on 5 clusters are presented in Appendix Table C.4. The basic characteristics of the original 4 clusters and the labels we attach to them remain similar. When given the opportunity to isolate a 5th cluster, the k-means procedure creates a cluster of skill-intensive industries.[18] The income regression results with 5 clusters are presented in Appendix Figure C.5. The 5th cluster itself does not have a positive impact on income, indeed both the foreign demand and foreign supply coefficients are relatively precisely estimated zeros. The main findings regarding the income impacts of the other clusters are preserved.

[17] We follow Belloni et al. (2014a) and choose the tuning parameter for the double-LASSO procedure through K-fold cross validation: see Appendix C.2.4. The statistics literature often chooses the tuning parameter to be one standard deviation above the minimizing value in order to select a more parsimonious model. Our baseline specification uses the minimizing value, which results in more controls being selected. We also check robustness to using a smaller tuning parameter for different specifications in Appendix Figures C.4 and C.9.

[18] The mean skilled labor share of this cluster, 0.54, is 21 percentage points higher than the skilled labor share of the second-most skill-intensive cluster.

In the second cluster robustness exercise, we assess the importance of sectors at the margins of the cluster classification. We add noise (standard deviation of 10% of the actual variability) to each characteristic of each sector, re-cluster sectors, and perform the full double-LASSO estimation using the new clusters. We repeat this procedure 1000 times. The goal of this procedure is to see how the cluster-specific income-impact coefficients are affected by switching a small number of marginal sectors from one cluster to another.

Appendix Figure C.6 reports the results. The dots indicate our baseline coefficient estimates, whereas the dashed bars indicate the 95% range of outcomes across simulations (not confidence intervals). For foreign demand shocks, the figure reveals that many of the coefficient estimates are quite stable: the range of estimates across simulations for raw materials processing, agriculture, and mining clusters is very small. On the other hand, reclassification tends to boost the coefficients on complex intermediate and consumption goods at the expense of the coefficient on capital goods. These results indicate that our most robust findings are that foreign demand shocks in raw materials processing, agriculture and mining have small income impacts and those in complex intermediates have large income impacts, while the results for the other sectors are less robust. In contrast, the results regarding the foreign supply shocks are essentially all quite fragile. Combined with the large confidence intervals in the baseline, this exercise indicates that the data yield very little useful information regarding the impact of foreign supply shocks.

Dropping Large and Contiguous Trading Partners

We next assess the sensitivity of the results to possible violations of the small country assumption. Country i can be a large trading partner of country n, such that the fixed effects estimated for country n are affected by the shocks to country i itself. Note that this concern is mitigated by the fact that the fixed effects are extracted from the gravity equations using the leave-one-out approach, whereby country i is dropped from the gravity sample when estimating the fixed effects that go into building country i's FMA's and CMA's. Nonetheless, we check the robustness of the results by dropping the countries for whom i is a large trading partner from the computation of the market access terms.

Specifically, when constructing the country i's FMA in sector k, we drop importer n from the summation in equation (3.22) if more than 25% of its imports in sector k are from country i, i.e. $\lambda_{ink,t} > 0.25$. When constructing the country n's CMA in sector k, we drop exporter i from the summation in equation (3.23) if more than 25%

of its exports in sector k go to country n, i.e. $\lambda_{ink,t} > 0.25$. The results are reported in Appendix Figure C.7. The results are broadly similar to the baseline, especially for the more robust results on the demand side.

Our identification relies on the assumption that country i's unobserved productivity shocks are uncorrelated with the foreign market access regressors. This assumption could be violated if productivity shocks are spatially correlated, so that nearby countries are subject to similar productivity shocks. To address this concern, we omit contiguous countries from the calculation of the market access terms and re-estimate the model. The results are reported in Appendix Figure C.8, and reveal very little change relative to the baseline.

3.5.3 Developed vs. Developing Countries

Our main specification pools all countries and time periods together and clusters on the industry dimension alone. It is also interesting to consider clustering along the country dimension, i.e. whether the impact of foreign shocks exhibits heterogeneity across different groups of countries.[19] One of the more intriguing possibilities is that rich and poor countries systematically differ in the income impact of foreign shocks to different sectors. To investigate this hypothesis, we split the sample into two groups based on the World Bank's 2016 country classification by income. Developing countries are those assigned by the World Bank to "low income" and "lower middle income" categories, and the developed countries the remaining group. According to this classification, 70 countries belong to the developed group, and 57 to the developing group. We then estimate elasticities of real income with respect to foreign shocks for the two country groups separately.

Figure 3.2 reports the results of the baseline specifications for the developed and developing groups. For both groups, the coefficients on demand shocks in complex intermediates are positive and statistically significant, although the magnitude is larger for the developed country groups. On the other hand, the capital goods coefficients behave very differently in the two samples: it is similar to the baseline coefficient in the developed country sample, but is very large and relatively precisely estimated in the developing country sample. For foreign supply shocks, developed countries exhibit high income impacts from consumer goods only, while developing countries additionally see high (and precisely estimated) income impacts from capital goods

[19] This heterogeneity could come from a combination of differences in underlying parameter values and in the point of approximation.

imports. The large coefficient on the capital goods supply shock is consistent with an earlier literature that documented strong correlations between prices/quantities of capital goods (imports or domestic) and economic development in poor countries (De Long and Summers, 1991; Lee, 1995; Eaton and Kortum, 2001b; Caselli and Wilson, 2004). One might expect that this finding would imply that foreign demand shocks in capital goods would be worse for poor countries, since they would tend to raise the domestic price of these goods. The fact that they do not suggests that poor countries experience significant increasing returns to the production of these goods.

We repeat each of the robustness checks described above for the rich and poor country sample split, with the results reported in Appendix Figures C.9-C.13. The main results are robust to these different specifications. Interestingly, the measurement error simulation for the split sample indicates much more stability across simulations that the baseline case, at least for foreign demand shocks.

3.6 Quantitative Implications

To assess the economic significance of the estimated coefficients, we perform two counterfactual exercises. The first is designed to illustrate the role of comparative advantage. Above, we found that foreign shocks in certain sectors have a higher income impact than in others. As a result, even a foreign shock that is completely uniform across sectors would be predicted to change real income differently across countries, depending on their initial trade shares. To get a sense of the extent of this heterogeneity, we compute the elasticity of each country's income to a worldwide uniform log-change in FMA and CMA, that is the same in every foreign sector and every foreign country. A simple transformation of our estimating equation leads to the following expression for this elasticity:

$$\frac{d \ln y_{i,t}}{d \ln FMA} = \sum_{g \in G} \widehat{\delta}_g^{ex} \sum_{k \in G} \lambda_{ik,t}^{ex},$$

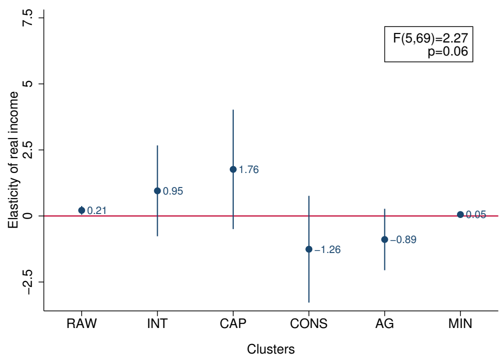
and

$$\frac{d \ln y_{i,t}}{d \ln CMA} = \sum_{g \in G} \widehat{\delta}_g^{im} \sum_{k \in G} \lambda_{ik,t}^{im}.$$

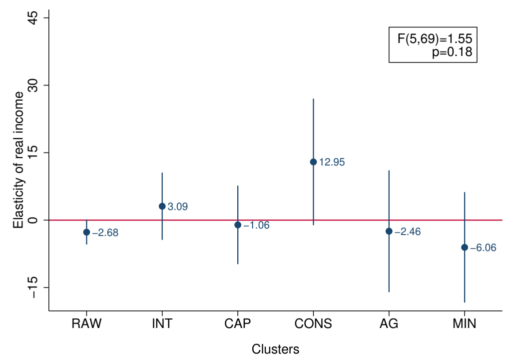
By imposing uniform foreign shocks across all countries and sectors, this counterfactual allows us to focus purely on the role of industrial specialization, as reflected in the $\lambda_{ik,t}$'s. Countries that have high export shares in clusters with a high estimated income

Figure 3.2 Developed vs. Developing Countries: Cluster-Specific Coefficients and Confidence Intervals

I. Developed Countries

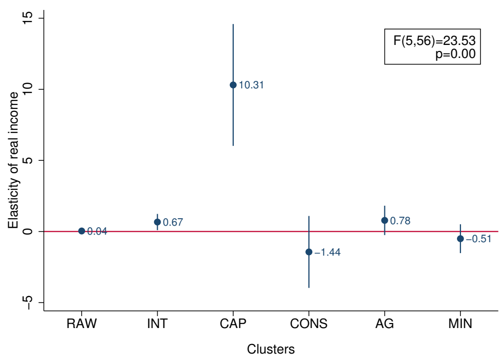


(a) Foreign Demand Shocks

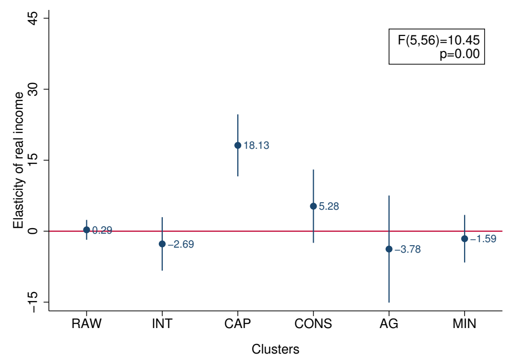


(b) Foreign Supply Shocks

II. Developing Countries



(c) Foreign Demand Shocks



(d) Foreign Supply Shocks

Notes: This figure reports the coefficients in estimating Equation (3.13), for the foreign demand shocks (FMA) (left panels), and foreign supply shocks (CMA) (right panels). The top panel displays the results for the sample of developed countries. 16 control variables are selected in the double-selection step. The bottom panel displays the results for developing countries. 0 control variables are selected in the double-selection step. The bars display the 95% confidence bands, that use standard errors clustered by country. The specifications control for initial GDP per capita. The boxes display the results of an F-test for equality of the coefficients in each plot.

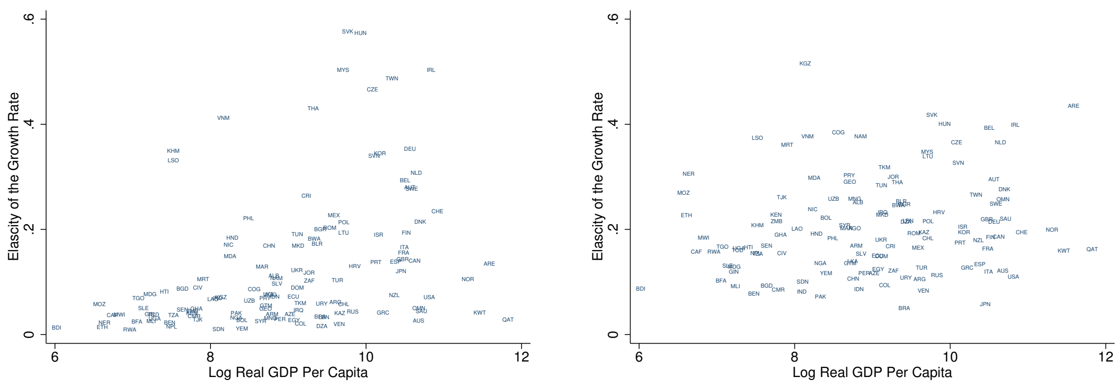
impact will have a more positive real income response.

The resulting elasticities calculated based on the 2015 import and export shares and the double-LASSO estimates from the bottom panel of Figure 3.1 are plotted in Figure 3.3 against log PPP-adjusted income per capita.[20] There is indeed a great deal of heterogeneity in the country impact of foreign shocks. The income elasticity

[20] As a robustness check, Appendix Figure C.14 plots the same elasticities using the estimates from Figure 3.2, which vary across countries according to income. Despite large differences in the estimates for capital goods, the resulting export demand elasticities with respect to the uniform shock are quite similar for most countries. This is because while the capital goods foreign demand shocks have a large coefficients among developing countries, capital goods exports are quantitatively small for most poor countries (2.5% of exports on average).

with respect to foreign demand shocks (left panel) ranges from essentially zero for countries chiefly in Sub-Saharan Africa, to 0.4-0.5 for some Central European and East Asian countries such as Hungary, Slovakia, Malaysia, and Taiwan. There is a similar level, and a similar amount of heterogeneity in the elasticity of income with respect to foreign supply shocks (right panel). Here, the relationship with per capita income is not apparent, as countries in virtually all income groups experiencing about the same range. Given the large differences in the estimated coefficients, this relative uniformity across countries indicates that the cluster-level import shares do not vary much across countries, and not systematically with per capita income.

Figure 3.3 Elasticity of Real Income with Respect to Foreign Shocks



(a) Foreign Demand Shocks (b) Foreign Supply Shocks

Notes: This figure presents the scatterplot of elasticity of income rate with respect to the foreign demand shocks (FMA) (left panel), and foreign supply shocks (CMA) (right panel) against real GDP per capita. Elasticity of income is calculated using the baseline estimates of coefficients in estimating equation (3.13) and the sectoral export and import shares in 2015.

Having illustrated the impact of heterogeneity in countries' comparative advantage, our next counterfactual is designed to illustrate the role of geography. Even though there is only one importer fixed effect for each country in each sector, the same vector of worldwide importer effects is experienced differently by each exporter due to its geographic position. As an example, there is only one change in the demand for capital goods in Germany, and one in China. Suppose that in a particular period, the importer effects reveal that China is having a much larger demand shock for capital goods than does Germany. This pair of importer-specific shocks will affect Belgium and Vietnam quite differently, as Vietnam is closer to China than to Germany, and the opposite is true for Belgium. What we would like to understand is how large is this type of heterogeneity. We thus construct counterfactual real income changes that would occur if Belgium experienced Vietnam's market access shocks. This counterfactual answers the question: how much would Belgium's real income change if in a particular time period it were picked up and moved to the place on the globe occupied by Vietnam? We do this for every pair of countries and in each decade.

Table 3.2 Predicted Annual Growth Difference, 2005-2015

	Growth difference, actual vs:		
	Median	25th pctile	75th pctile
G7			
Canada	-1.55	-1.95	-1.07
France	-0.89	-1.18	-0.51
Germany	-1.31	-1.68	-0.76
Italy	-0.56	-0.75	-0.29
Japan	0.80	0.66	0.96
UK	-1.43	-1.75	-1.04
US	0.02	-0.13	0.20
BRICS			
Brazil	0.03	-0.06	0.16
China	-1.63	-1.86	-1.26
India	0.37	0.21	0.48
Russia	-0.03	-0.26	0.33
South Africa	0.32	-0.05	0.69

Notes: This table reports the differences in real income growth, in percent per annum, between the actual growth and the counterfactual growth that the country would experience if it were moved to the median (resp. 25th and 75th percentile) geographic position.

To begin getting a sense of the magnitudes involved, Table 3.2 reports the results for a set of prominent countries, namely the G7 and the BRICS. The first column reports the difference between the country's actual growth and the growth that would obtain if the country were moved to the position of the median country, where "median" means the median difference among all the possible counterfactual geographic positions. So, a value of 1 in the first column implies that the country grew 1 percentage point per annum faster in its actual geographic position, relative to being moved to the median position in the world. The second and third columns report the counterfactual growth differences due to being moved to the 25th and the 75th percentile geographic position for that country.

A few features of the table stand out. First, the numbers are large and heterogeneous. In this period, most of the G7 countries actually grew substantially *slower* than they could have in an alternative geographic position, and some of these growth differentials are substantial, between 0.5 and 1.5 percent annually. The exception to this pattern is Japan, which grew 0.8 percentage points faster than it would have in the median geographic position. The picture for the BRICS is less clear, with medians closer to zero. The exception is China, which would have been better off locating in the median position.

Table 3.3 reports the summary statistics by region and period. The two regions at the extremes are East Asia & Pacific and Western Europe/North America. The median country in East Asia has reaped a substantial and increasing benefit of geographic location. In the most recent decade, its growth has been 0.8 percentage points per annum higher than it would have been had it been located at the median geographic location in the world. This benefit of East Asian location has been consistently positive across 5 decades, and if anything increasing over time. On the opposite end, the typical country in the Western Europe/North America region has for the most part grown slower than it would have had it been moved to the median location. This may first appear surprising, as these are some of the richest and most open countries in the world. However, these comparisons capture the impact of *changes* in foreign demand on economic growth rates. So the negative growth differentials are perfectly consistent with West European countries having high market access *levels*. What these results reveal is that these wealthy countries are located next to relatively slow-growing countries, and thus foreign demand and supply have expanded more slowly for them than they would have if they had been located in faster-growing regions of the world.

In other groups of countries, the overall growth impact of geographic location is quite a bit smaller overall, and switches sign over time. The absolute impact of geography on growth tends to rise over time, as countries become more open overall. In the last decade, the Middle East, South Asia, and Sub-Saharan Africa have enjoyed a modest benefit of their geographic position, whereas for Latin America and Eastern Europe/Central Asia, their location has had a modest cost.

Table 3.3 Predicted Annual Growth Difference Relative to Median Geographic Location, Medians by Region and Time Period

Region	1975	1985	1995	2005	2015
East Asia & Pacific	0.44	0.53	0.08	0.75	0.78
	[0.07 , 0.61]	*[0.25 , 0.71]*	*[-0.13 , 0.27]*	*[0.36 , 2.08]*	*[0.29 , 1.83]*
	10	14	14	14	14
Eastern Europe & Central Asia	0.00	-0.36	0.11	-0.06	-0.19
	[-0.01 , 0.00]	*[-0.64 , -0.16]*	*[0.08 , 0.40]*	*[-0.24 , 0.31]*	*[-0.82 , 0.04]*
	2	6	6	24	24
Latin America & Caribbean	-0.28	-0.12	-0.01	-0.06	-0.14
	[-0.44 , -0.02]	*[-0.27 , 0.03]*	*[-0.23 , 0.23]*	*[-0.14 , 0.08]*	*[-0.53 , -0.12]*
	18	18	18	18	18
Middle East & North Africa	0.03	-0.06	-0.03	0.01	0.17
	[-0.04 , 0.05]	*[-0.65 , 0.15]*	*[-0.13 , 0.14]*	*[-0.23 , 0.17]*	*[-0.08 , 0.46]*
	7	14	14	15	15
South Asia	0.05	0.07	0.06	0.03	0.34
	[-0.01 , 0.15]	*[0.04 , 0.21]*	*[0.01 , 0.11]*	*[-0.02 , 0.19]*	*[0.34 , 0.36]*
	4	5	5	5	5
Sub-Saharan Africa	-0.03	0.07	-0.23	0.12	0.22
	[-0.21 , 0.08]	*[-0.05 , 0.24]*	*[-0.44 , -0.11]*	*[0.06 , 0.35]*	*[0.00 , 0.31]*
	28	30	30	30	33
West Europe/North America	0.01	-0.88	0.57	-0.37	-0.87
	[-0.21 , 0.09]	*[-1.32 , -0.52]*	*[0.31 , 1.01]*	*[-1.04 , -0.09]*	*[-1.61 , -0.63]*
	18	18	18	18	18

Notes: This table reports the region- and period-specific differences in economic growth, in percent per annum, between the actual growth and the counterfactual growth that the country would experience if it were moved to the median geographic position. The numbers in square brackets are the interquartile range across countries in the region and time period. The bottom rows report the number of countries in each cell.

3.7 Conclusion

Using a theoretically grounded approach and employing new empirical techniques, we have shown that there is significant heterogeneity in the impact of foreign shocks in different sectors. Positive foreign demand shocks in sectors producing complex intermediate and capital goods have a significantly higher real income impact than shocks in other sectors, while positive supply shocks to capital goods are especially beneficial to developing countries. Our quantitative results imply that the interaction between initial comparative advantage and the pattern of foreign shocks is important for understanding the variety of growth experiences across countries.

Our findings do not have immediate implications for policy, except perhaps that countries should pursue increased market access more vigorously in some sectors relative to others. However, questions surrounding the effect of the external environment on economic development for developing countries have been central in the great policy debates of the past 60 years, from import-substituting industrialization to the Washington Consensus to the "Washington Confusion" (Rodrik, 2006). Our results speak to these debates insofar as they affirm the importance of the external environment for real income and validate a focus on the sectoral dimensions of policy. A fuller understanding of optimal sectoral policy requires considering domestic policies as well (Bartelme et al., 2019), along with the ever-mysterious drivers of productivity growth.

Table 3.4 Predicted Annual Growth Difference Relative to Median Geographic Location, Medians by Region and Time Period

Actual Region	Counterfactual Region: East Asia & Pacific	Eastern Europe & Central Asia	Latin America & Caribbean	Middle East & North Africa	South Asia	Sub-Saharan Africa	West Europe/ North America
East Asia & Pacific (N = 14)	0.07 *[-0.09 , 0.63]*	0.87 *[0.53 , 1.92]*	1.05 *[0.55 , 2.19]*	0.52 *[0.12 , 1.45]*	-0.06 *[-0.19 , 0.54]*	0.60 *[0.09 , 1.69]*	1.24 *[1.02 , 2.97]*
Eastern Europe & Central Asia (N = 24)	-1.06 *[-2.26 , -0.60]*	0.08 *[-0.48 , 0.40]*	0.09 *[-0.38 , 0.63]*	-0.32 *[-0.92 , 0.05]*	-1.12 *[-2.26 , -0.68]*	-0.40 *[-1.16 , -0.20]*	0.76 *[0.37 , 1.26]*
Latin America & Caribbean (N = 18)	-0.56 *[-1.02 , -0.42]*	-0.04 *[-0.26 , 0.01]*	-0.01 *[-0.17 , 0.09]*	-0.27 *[-0.63 , -0.17]*	-0.67 *[-1.13 , -0.51]*	-0.35 *[-0.64 , -0.20]*	0.35 *[0.17 , 0.55]*
Middle East & North Africa (N = 15)	-0.41 *[-1.03 , 0.00]*	0.46 *[-0.05 , 0.66]*	0.45 *[0.16 , 0.76]*	0.15 *[-0.29 , 0.34]*	-0.51 *[-0.96 , -0.09]*	0.07 *[-0.44 , 0.17]*	1.10 *[0.18 , 1.37]*
South Asia (N = 5)	0.08 *[0.03 , 0.14]*	0.37 *[0.36 , 0.40]*	0.52 *[0.39 , 0.55]*	0.29 *[0.28 , 0.38]*	0.08 *[0.04 , 0.13]*	0.28 *[0.28 , 0.29]*	0.64 *[0.47 , 0.75]*
Sub-Saharan & Africa (N = 33)	-0.31 *[-0.64 , -0.14]*	0.36 *[0.09 , 0.64]*	0.37 *[0.12 , 0.58]*	0.02 *[-0.07 , 0.14]*	-0.47 *[-0.70 , -0.29]*	0.07 *[-0.12 , 0.14]*	0.98 *[0.70 , 1.22]*
West Europe/ North America (N = 18)	-1.46 *[-2.31 , -1.21]*	-0.61 *[-1.18 , -0.30]*	-0.72 *[-1.31 , -0.42]*	-1.00 *[-1.69 , -0.70]*	-1.67 *[-2.58 , -1.40]*	-1.08 *[-1.87 , -0.91]*	-0.03 *[-0.56 , 0.52]*

Notes: This table reports the region pair-specific differences in economic growth, in percent per annum, between the actual growth and the counterfactual growth that the country would experience if it were moved to the counterfactual region. The numbers in square brackets are the interquartile range across countries in the region.

Appendices

Appendix A

Appendices of Chapter I

A.1 Data and Estimation Appendix

A.1.1 Census Datasets

Technical notes on the data construction is under view and will be disclosed soon.

A.1.2 EG coagglomeration index

Ellison and Glaeser (1997) propose a measure of coagglomeration of a group of I activities. Denote ω^i as the activity i's employment share in the I activities. Let $s_1, \ldots, s_L$ be the share of total employment of the I activities in each of the geographic subareas, where $s_l = \sum_i \omega^i s_l^i$. Let $G = \sum_l (s_l - x_l)^2$ be the raw geographic concentration for the I-activity group and $H = \sum_i (\omega^i)^2 H_i$ be the Herfindahl index of the I-activity group. The EG index of coagglomeration is

$$\gamma^c = \frac{[G/(1 - \sum_l x_l^2)] - H - \sum_i \gamma^i (\omega^i)^2 (1 - H^i)}{1 - \sum_i (\omega^i)^2} \tag{A.1}$$

where γ^i measures the agglomeration of activity i. The coagglomeration index reflects excess geographic concentration of the I-activity group relative to what would be expected if each activity were as agglomerated as it is when the locations of the agglomeration were independent.

The coagglomeration index takes on a simpler form when used to measure coagglomeration of two activities.

$$\gamma^c = \frac{\sum_l (s_l^a - x_l)(s_l^b - x_l)}{1 - \sum_l x_l^2} \tag{A.2}$$

Proof The raw geographic concentration for the 2-activity group can be written as

$$\begin{aligned} G &= \sum_l \left(\omega^a s_l^a + \omega^b s_l^a - \left(\omega^a + \omega^b\right) x_m\right)^2 \\ &= \omega^a G^a + \omega^b G^b + 2\omega^a\omega^b \sum_l (s_l^a - x_l)(s_l^b - x_l) \end{aligned}$$

The aggomeration index of activity i is

$$\gamma^i \equiv \frac{\sum_l \left(s_l^i - x_l\right)^2 - \left(1 - \sum_l x_l^2\right) H^i}{\left(1 - \sum_l x_l^2\right)\left(1 - H^i\right)} \tag{A.3}$$

Plug in γ_i and G into Equation A.1

$$\begin{aligned} \left(1 - \sum_i \left(\omega^i\right)^2\right)\gamma^c &= \left[G/\left(1 - \sum_l x_l^2\right)\right] - H - \sum_i \gamma^i \left(\omega^i\right)^2 \left(1 - H^i\right) \\ &= \frac{G - H\left(1 - \sum_l x_l^2\right) - \sum_i \left(\omega^i\right)^2 \left(\sum_l \left(s_l^i - x_l\right)^2 - \left(1 - \sum_l x_l^2\right) H^i\right)}{1 - \sum_l x_l^2} \\ &= \frac{\omega^a G^a + \omega^b G^b + 2\omega^a\omega^b \sum_l (s_l^a - x_l)(s_l^b - x_l) - \omega^a G^a - \omega^b G^b}{1 - \sum_l x_l^2} \\ &= \frac{\left(1 - \sum_i \left(\omega^i\right)^2\right) \sum_l (s_l^a - x_l)(s_l^b - x_l)}{1 - \sum_l x_l^2} \end{aligned}$$

Thus, we have

$$\gamma^c = \frac{\sum_l (s_l^a - x_l)(s_l^b - x_l)}{1 - \sum_l x_l^2} \tag{A.4}$$

Table A.1 The coagglomeration of cross-industry production VS. the coagglomeration of within-industry innovation and production

NAICS	NAICS Industry Description	Cross-indutry Pairwise Production					Within-industry Innovation and Production
		Mean	25th percentile	Median	75th percentile	Max	
3111	Animal Food Manufacturing	0.0006	-0.0001	0.0003	0.0010	0.0130	0.0009
3112	Grain and Oilseed Milling	0.0003	0.0000	0.0006	0.0010	0.0018	0.0095
3113	Sugar and Confectionery Product Manufacturing	0.0004	0.0001	0.0004	0.0008	0.0032	0.0162
3114	Fruit and Vegetable Preserving and Specialty Food Manufacturing	0.0004	0.0002	0.0004	0.0006	0.0012	0.0042
3115	Dairy Product Manufacturing	0.0004	0.0001	0.0004	0.0006	0.0013	-0.0029
3116	Animal Slaughtering and Processing	0.0004	0.0000	0.0006	0.0010	0.0025	0.004
3117	Seafood Product Preparation and Packaging	0.0002	-0.0007	-0.0002	0.0006	0.0081	0.0277
3118	Bakeries and Tortilla Manufacturing	0.0003	-0.0004	0.0001	0.0007	0.0089	0.0016
3119	Other Food Manufacturing	0.0004	-0.0001	0.0003	0.0007	0.0077	0.0059
3121	Beverage Manufacturing	0.0001	0.0000	0.0001	0.0003	0.0010	-0.0002
3131	Fiber, Yarn, and Thread Mills	0.0006	-0.0003	0.0006	0.0014	0.0114	0.0127
3132	Fabric Mills	0.0006	0.0001	0.0004	0.0010	0.0056	0.0036
3133	Textile and Fabric Finishing and Fabric Coating Mills	0.0012	-0.0012	0.0001	0.0021	0.0457	0.0152
3141	Textile Furnishings Mills	0.0008	0.0001	0.0004	0.0008	0.0114	0.0194
3149	Other Textile Product Mills	0.0003	-0.0002	0.0000	0.0003	0.0048	0.0006
3151	Apparel Knitting Mills	0.0008	-0.0001	0.0004	0.0010	0.0115	0.0102
3152	Cut and Sew Apparel Manufacturing	0.0008	-0.0065	-0.0018	0.0054	0.0457	0.0463
3159	Apparel Accessories and Other Apparel Manufacturing	0.0003	-0.0019	-0.0006	0.0010	0.0384	0.0105
3161	Leather and Hide Tanning and Finishing	0.0006	-0.0002	0.0005	0.0014	0.0049	0.035
3162	Footwear Manufacturing	0.0002	-0.0006	-0.0001	0.0005	0.0129	0.0188
3169	Other Leather and Allied Product Manufacturing	0.0010	-0.0015	-0.0001	0.0020	0.0448	0.0189
3211	Sawmills and Wood Preservation	0.0002	-0.0003	0.0004	0.0012	0.0032	0.0051
3212	Veneer, Plywood, and Engineered Wood Product Manufacturing	0.0001	-0.0004	0.0003	0.0009	0.0032	0.0039
3219	Other Wood Product Manufacturing	0.0003	-0.0001	0.0004	0.0008	0.0040	0.0053
3221	Pulp, Paper, and Paperboard Mills	0.0003	-0.0003	0.0005	0.0012	0.0032	0.0068
3222	Converted Paper Product Manufacturing	0.0005	0.0002	0.0006	0.0008	0.0019	0.0002
3231	Printing and Related Support Activities	0.0004	0.0000	0.0004	0.0007	0.0033	0.0009
3241	Petroleum and Coal Products Manufacturing	0.0004	-0.0002	0.0001	0.0006	0.0079	0.0196

3251	Basic Chemical Manufacturing	0.0002	-0.0005	0.0002	0.0011	0.0071	0.0099
3252	Resin, Synthetic Rubber, and Artificial Synthetic Fibers and Filaments Manufacturing	0.0004	-0.0002	0.0006	0.0010	0.0043	0.0119
3253	Pesticide, Fertilizer, and Other Agricultural Chemical Manufacturing	0.0001	-0.0002	0.0003	0.0009	0.0018	0.0078
3254	Pharmaceutical and Medicine Manufacturing	0.0001	-0.0003	-0.0001	0.0004	0.0020	0.0177
3255	Paint, Coating, and Adhesive Manufacturing	0.0006	-0.0001	0.0005	0.0012	0.0030	0.0043
3256	Soap, Cleaning Compound, and Toilet Preparation Manufacturing	0.0005	-0.0003	0.0004	0.0011	0.0114	0.0072
3259	Other Chemical Product and Preparation Manufacturing	0.0005	0.0002	0.0004	0.0008	0.0014	0.0109
3261	Plastics Product Manufacturing	0.0005	0.0002	0.0006	0.0008	0.0018	0.0013
3262	Rubber Product Manufacturing	0.0005	0.0001	0.0005	0.0009	0.0021	0.0061
3271	Clay Product and Refractory Manufacturing	0.0003	0.0000	0.0004	0.0009	0.0025	0.0069
3272	Glass and Glass Product Manufacturing	0.0004	0.0001	0.0003	0.0007	0.0015	0.0069
3273	Cement and Concrete Product Manufacturing	-0.0001	-0.0004	0.0000	0.0005	0.0014	0.0005
3274	Lime and Gypsum Product Manufacturing	0.0000	-0.0003	0.0001	0.0005	0.0018	0.0025
3279	Other Nonmetallic Mineral Product Manufacturing	0.0002	0.0000	0.0004	0.0006	0.0015	0.0053
3311	Iron and Steel Mills and Ferroalloy Manufacturing	0.0003	-0.0006	0.0006	0.0013	0.0046	0.0179
3312	Steel Product Manufacturing from Purchased Steel	0.0006	0.0000	0.0006	0.0012	0.0038	0.0029
3313	Alumina and Aluminum Production and Processing	0.0004	0.0000	0.0005	0.0009	0.0027	0.009
3314	Nonferrous Metal (except Aluminum) Production and Processing	0.0005	0.0002	0.0005	0.0009	0.0027	0.0082
3315	Foundries	0.0007	0.0002	0.0007	0.0011	0.0023	0.0049
3321	Forging and Stamping	0.0007	0.0000	0.0007	0.0013	0.0032	0.0057
3322	Cutlery and Handtool Manufacturing	0.0004	0.0002	0.0005	0.0009	0.0019	0.0046
3323	Architectural and Structural Metals Manufacturing	0.0002	0.0000	0.0002	0.0004	0.0023	0.0015
3324	Boiler, Tank, and Shipping Container Manufacturing	0.0005	0.0002	0.0006	0.0011	0.0027	0.0071
3325	Hardware Manufacturing	0.0007	0.0002	0.0006	0.0011	0.0060	0.0131
3326	Spring and Wire Product Manufacturing	0.0005	0.0000	0.0005	0.0010	0.0034	0.013
3327	Machine Shops; Turned Product; and Screw, Nut, and Bolt Manufacturing	0.0007	0.0001	0.0005	0.0011	0.0039	0.0022
3328	Coating, Engraving, Heat Treating, and Allied Activities	0.0008	0.0000	0.0008	0.0015	0.0080	0.0063
3329	Other Fabricated Metal Product Manufacturing	0.0004	0.0000	0.0004	0.0008	0.0029	0.0018
3331	Agriculture, Construction, and Mining Machinery Manufacturing	0.0002	-0.0005	0.0002	0.0013	0.0071	0.0142
3332	Industrial Machinery Manufacturing	0.0005	0.0001	0.0006	0.0009	0.0033	0.0074
3333	Commercial and Service Industry Machinery Manufacturing	0.0004	-0.0001	0.0004	0.0009	0.0025	0.0029

3334	Ventilation, Heating, Air-Conditioning, and Commercial Refrigeration Equipment Manufacturing	0.0003	0.0001	0.0005	0.0009	0.0021	0.0031
3335	Metalworking Machinery Manufacturing	0.0006	0.0000	0.0005	0.0012	0.0047	0.0111
3336	Engine, Turbine, and Power Transmission Equipment Manufacturing	0.0004	-0.0001	0.0005	0.0012	0.0028	0.0179
3339	Other General Purpose Machinery Manufacturing	0.0005	0.0002	0.0006	0.0010	0.0019	0.0037
3341	Computer and Peripheral Equipment Manufacturing	0.0001	-0.0007	0.0000	0.0009	0.0050	0.0116
3342	Communications Equipment Manufacturing	0.0007	-0.0018	-0.0003	0.0016	0.0281	0.0199
3343	Audio and Video Equipment Manufacturing	0.0007	-0.0007	0.0002	0.0010	0.0165	0.0199
3344	Semiconductor and Other Electronic Component Manufacturing	0.0003	-0.0007	-0.0002	0.0003	0.0099	0.0212
3345	Navigational, Measuring, Electromedical, and Control Instruments Manufacturing	0.0005	-0.0005	0.0000	0.0008	0.0063	0.0088
3346	Manufacturing and Reproducing Magnetic and Optical Media	0.0007	-0.0006	0.0000	0.0010	0.0133	0.0422
3351	Electric Lighting Equipment Manufacturing	0.0006	-0.0006	0.0003	0.0011	0.0178	0.0031
3352	Household Appliance Manufacturing	0.0004	0.0000	0.0009	0.0013	0.0021	0.0117
3353	Electrical Equipment Manufacturing	0.0005	0.0000	0.0005	0.0010	0.0034	0.0032
3359	Other Electrical Equipment and Component Manufacturing	0.0006	0.0003	0.0005	0.0008	0.0017	0.0034
3361	Motor Vehicle Manufacturing	0.0003	-0.0003	0.0005	0.0012	0.0076	0.0498
3362	Motor Vehicle Body and Trailer Manufacturing	0.0006	-0.0002	0.0007	0.0013	0.0040	0.0283
3363	Motor Vehicle Parts Manufacturing	0.0005	-0.0001	0.0006	0.0011	0.0076	0.0185
3364	Aerospace Product and Parts Manufacturing	0.0003	-0.0011	-0.0004	0.0007	0.0193	0.0248
3365	Railroad Rolling Stock Manufacturing	0.0003	-0.0002	0.0005	0.0012	0.0049	0.0151
3366	Ship and Boat Building	-0.0001	-0.0007	0.0001	0.0006	0.0023	0.0389
3369	Other Transportation Equipment Manufacturing	0.0006	0.0001	0.0006	0.0010	0.0031	0.0114
3371	Household and Institutional Furniture and Kitchen Cabinet Manufacturing	0.0006	0.0001	0.0003	0.0008	0.0067	0.0038
3372	Office Furniture (including Fixtures) Manufacturing	0.0005	0.0002	0.0004	0.0007	0.0022	0.0123
3379	Other Furniture Related Product Manufacturing	0.0002	-0.0005	-0.0001	0.0006	0.0102	0.0042
3391	Medical Equipment and Supplies Manufacturing	0.0003	-0.0002	0.0001	0.0006	0.0036	0.0092
3399	Other Miscellaneous Manufacturing	0.0002	-0.0002	0.0002	0.0005	0.0041	0.0032

Notes: Column 3-7 reports the mean, 25th percentile, median, 75th percentile and max coagglomeration of cross-industry production pairs, respectively. The last column reports the within-industry coagglomeration of innovation and production.

Appendix B
Appendices of Chapter II

B.1 Data Appendix

B.1.1 Constructing Percentile-Level Expenditure Weights

Consumer Expenditure Survey

We use data from the Consumer Expenditure Survey (CES) to obtain the expenditure weights of consumers. The CES data are collected by the Census Bureau, and cover expenditures, income, and demographic characteristics of households in the United States. The CES is the primary source of data for constructing the weights for the US Consumer Price Index.

The CES contains two modules, the Diary and the Interview. The Diary is designed to measure expenditures on daily items, such as groceries, personal products, and other frequent purchases. The Interview is designed to measure large or durable expenditures, such as major appliances, vehicles, and other large infrequent purchases. The Diary records household spending for two consecutive survey reference weeks, while the Interview records purchases over the previous three months.

For each survey, we make use of expenditure, income, and characteristics files in computing expenditure weights. In the expenditure files, the CES collects household expenditures on about 600 Universal Classification Code (UCC) categories. Questions such as "How much did you spend on babysitting in the last quarter" are asked in the survey and the corresponding responses are saved in *UCC 340210 babysitting and child care*. Overall, there are questions on about 350 UCC categories in the Interview module, and on 250 UCCs in the Diary module. Income files record detailed information on household monthly income from different sources, such as wages and salaries, or interest and dividends. Characteristics files record demographic characteristics data for each member of the household, such as education, gender, race, etc. Income variables, which contain annual values for the 12 months prior to the interview month, are also included in the characteristics files.

Diary and Interview modules survey different households each year, so a household in the Diary will not appear in the Interview and vice versa. Thus we could never observe the full consumption profiles of an actual household and we could not compute expenditure shares for an actual household. Rather, we aggregate households into percentiles and work with the percentile-level household expenditure shares.

Constructing the Concordance

The in-scope expenditures for CPI could be divided into 8 major groups, 70 expenditure classes, 211 item strata (item level) and 303 entry level items (ELI). CPI uses the item strata -- e.g. *SEFT04 Spices, seasonings, condiments, sauces* -- as the elementary level of its expenditure weights and price index calculation. Within each item stratum, one or more substrata are defined as ELIs, which are the ultimate sample units for products. For example, there are four ELIs under item *SEFT04*: *FT041 Salt and other seasonings and spices*, *FT042 Olives, pickles and relishes*, *FT043 Sauces and gravies* and *FT044 Other condiments*.

Using CES data to compute the item-level and ELI-level expenditure weights from CES, we need a concordance between the UCC categories, item strata codes and the ELIs. The concordance is constructed by following the BLS document *"CPI requirements for CE"* Appendix B. The CES collects household expenditures on about 600 Universal Classification Code (UCC) categories, which could be concorded to 303 ELIs. To combine the expenditure weights with the frequency of price adjustment data from Nakamura and Steinsson (2008), we look at a subsample of 265 ELIs. And we could further aggregate the 265 ELIs to 178 item strata.

Compiling the Expenditure, Income, and Characteristics Files

To obtain the expenditure shares at the detailed product category level for households at different percentiles of the income distribution, we take the following steps.

In the first step, we put together the quarterly expenditure, income, and characteristics files from the Interview survey. With the compiled interview data, for each household, we could observe its interviewed month and year, monthly expenditures on the UCC categories in the previous three months as well as annual income for the 12 months prior to the interview. One thing to note is that respondents are asked to report expenditures made since the first of the three months prior to the interview month. For example, if a household is interviewed in February of 2015, they are reporting expenditures for November and December of 2014, and January of 2015. Thus, to produce a 2014 annual estimate based on expenditures made in 2014 (calendar period), one needs to access five collection-quarter files, the first quarter of 2014 through the first quarter of 2015.

By the same token, we put together the expenditure, income, and characteristics files from the Diary survey. For each household in the Diary survey, we are able to observe its weekly expenditure on the detailed UCC categories and its annual income for the 12 months prior to the interview. Then we append the compiled Interview data file to the compiled Diary one to get the whole sample of UCCs.

Adjusting the Expenditure Values

In the second step, we make several adjustments to the collected expenditures in order to meet the BLS's requirements for the creation of CPI expenditure weights. The adjustments are made following the BLS document *"CPI Requirements of CE"*.

Housing

Two adjustments are made to housing categories.

- **Owners' Equivalent Rent of Primary Residence**

  UCC categories only collect the value of the house, its property taxes, real estate fees, and mortgage interests. Houses and other residential structures are capital goods and should not be considered as CPI items. Interest costs (such as mortgage interest), property taxes and most maintenance costs, are part of the cost of the capital good and are not consumption expenditures either. All of these are not useful in computing the expenditure weights for the item *Owners' equivalent rent of primary residence.*

  According to the BLS document *"How the CPI measures price change of Owners' equivalent rent of primary residence (OER) and Rent of primary residence (Rent)"*, the expenditure weight in the CPI market basket for *Owners' equivalent rent of primary residence (OER)* is based on the following question that the CES asks consumers who own their primary residence:

  "If someone were to rent your home today, how much do you think it would rent for monthly, unfurnished and without utilities?"

  CES collects the household responses to this question and saves them in the variable *RENTEQVX* in characteristics files. We construct an artificial UCC code "999999" to store the values of variable *RENTEQVX*, which provides the household expenditure on the owners' equivalent rent of primary residence.

- **Homeowner Insurance/Maintenance/Major Appliance**
  The BLS adjusts the expenditures on homeowner insurance, maintenance, and major appliances to separate the consumption components of those expenditures from the investment component. The BLS uses a factor of 0.43 to account for the consumption portion of a homeowner's total expenditure on these housing categories. The factor is based on the likelihood that renters will purchase these types of appliances and perform these types of home maintenance and improvement. Thus, to reflect the consumption portion of a homeowner's total expenditure on housing insurance, maintenance, and major appliances, we multiply the expenditures on the corresponding UCC categories by 0.43.

Medical Care

The BLS uses the National Health Expenditure (NHE) tables produced by the Center for Medicare and Medicaid Services (CMS) to calculate the factors that redistribute the weights from private health insurance and Medicare premium to medical care services. Unfortunately, we do not have access to the underlying formulas the BLS used to calculate these factors. By way of approximation, we take the redistributing factors from the NHE *Table 20 Private Health Insurance Benefits and Net Cost; Levels, Annual Percent Change and Percent Distribution, Selected Calendar Years 1960-2015.*[1]

We redistribute the expenditures from private health insurance and Medicare premiums related UCC categories to health care services categories, such as nursing homes and adult day services, by using factors obtained from the table mentioned above. Note that medical reimbursements are allocated across all households to smooth the household expenditures on medical expenses. That is to say, a household may be reimbursed even during a period in which they had no medical expenses.

Transportation

- **Used Cars**
  Expenditures on used cars and trucks should only reflect dealer value added.

[1]For more details see the link `https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/nationalhealthexpenddata/nhe-fact-sheet.html`

Therefore, the expenditure weights on used cars and trucks should be determined by spending on used cars and trucks, minus trade-in value of vehicles and other sales of consumer-owned vehicles.
CES does not provide data on trade-in values of vehicles (*UCC 450116* and *450216*) and other sales of consumer owned vehicles (*UCC 860100* and *860200*). Thus, we take the expenditure weight on used cars and trucks from the BLS released table *Relative Importance of Components in the Consumer Price Index* to recover the ratio of trade-in values and other sales of vehicles to spending on used cars and trucks, and we find the ratio is around 1/2. Thus, we reduce the spending on used cars and trucks to half to reflect only the dealer value added.

- **Gasoline**
  Gasoline expenditures are not allocated into categories (regular, premium, midgrade, etc.) at collection. To distribute the total gasoline expenditures (*UCC 470111*) amongst the gasoline ELIs (*TB011 Regular Unleaded Gasoline,TB012 Midgrade Unleased Gasoline and TB013 Premium Unleased Gasoline*), the BLS constructed the distribution factors from expenditure habits in each primary sampling unit (PSU).
  However, we don't have access to the expenditure habits of each PSU. Instead, we follow Nakamura and Steinsson (2008), and allocate the expenditures on gasoline to regular, premium, midgrade categories equally.

Aggregating Households into Percentiles

In the third step, we aggregate households into percentiles. Because the Interview and the Diary survey different households, we sort the households into percentiles in two sub-steps. First, we aggregate the households in the Interview survey into percentiles based on imputed household annual income before tax, and then find the income cut-offs for each percentile. Second, we use the Interview survey income cut-offs to divide households from the Diary survey into percentiles. In this case, each household in our data sample has been sorted into a percentile. We could get similar results by using income cutoffs from the Diary survey to aggregate households in the Interview survey into percentiles.

The CES data start to include the imputed income since 2004. Before that it only publishes income data collected from households that are complete income reporters. Households are defined as complete reporters if they report one of the major sources of income, such as wages and salaries, Social Security income, or self-employment income.

However, even a complete reporter might not provide information on all sources of income they indicate they received. Thus, in cases when the values of income are not reported, imputation allows them to be estimated. We sort households into percentiles based on the imputed household income before tax, which is only available since 2004. Because of this, therefore, we could only compute the percentile-level expenditure weights since 2004.

Table B.1 reports the income cutoffs and average incomes in the selected quantiles of the income distribution.

Table B.1 Income cutoffs and averages for selected quantiles of the income distribution in the CES

	Cut-offs			
	Lower	Upper	Median	Mean
Bottom 5%	-23,297	5,838	2,343	2,450
Middle 40-60%	36,504	62,808	48,828	48,969
96-99%	212,148	332,196	249,677	253,900
Top 1%	332,279	846,706	392,148	414,011

Notes: The table the range and the averages of the incomes in selected quantiles of the income distribution in the CES data.

Calculating the Expenditure Shares

In the final step, we calculate the expenditure shares at the detailed product category level for households at different income percentiles.

First, we calculate the average expenditure for each detailed UCC category for households at different income percentiles. Note that there is a distinction between survey period and expenditure reference period in the interview survey, as the CES collects household spending in the three months prior to the interview month. This distinction will affect the estimation procedure for producing household average expenditure during a calendar year. For example, households interviewed in February will report their spending for November and December of 2014 and January of 2015. Thus, to compute the average value for expenditures made on a certain UCC category during year 2015, they only contribute one month (January) of the expenditures they made during the expenditure reference period to the calculation. While households interviewed in May report their expense for February, March and April of 2015 and could contribute all their expenditures to compute the average expenditure this household made during 2015. To reflect the number of months a household can contribute to the mean value of a calendar year, we follow the BLS to create a variable called MO_SCOPE. In the above example, MO_SCOPE=1 for households interviewed in February and MO_SCOPE=3 for households interviewed in May. There is no such distinction between the survey period and expenditure reference period in the Diary. We multiply each weekly expenditure by 13 to get a corresponding quarterly expenditure. As there is no lag between the survey period and the expenditure reference period, the number of months households in the Diary survey contribute to estimate of the mean value is 3, i.e. MO_SCOPE=3. We could also interpret MO_SCOPE as the number of months a household reports expenditures during a calendar year.

Following the BLS manual, we use the formula below to calculate the average expenditure for each UCC category k at each percentile h. First, for household i at percentile h, we sum over all the spending it made on good k during the calendar year. Second, we weight total expenditures made by household i in percentile h on good k up by its household-specific sampling weight. Third, we sum up the weighted household expenditures on good k over all the households at percentile h. Fourth, we divide the sum of weighted household expenditures on good k at percentile h by the sum of the weighted number of months household at percentile h reported expenditures during the calendar year, to get the monthly average income on good k of household at percentile h. Then multiplying the monthly average expenditure by 12, we get the annualized average expenditure for each UCC category k at percentile h:

$$\bar{X}_k^h = \frac{\sum_i FINLWT_i^h \cdot \sum_t C_{i,k,t}^h}{\sum_i FINLWT_i^h \cdot MO\_SCOPE_i^h} \times 12$$

where $FINLWT_i^h$ is the sampling weight for household i at income percentile h, $C_{i,k,t}^h$ is the expenditure on good k of household i at income percentile h during month t, and $MO\_SCOPE_i^h$ denotes number of months household i reports expenditures during a calendar year.

Second, we take percentile-level average expenditure for each UCC from above, and then aggregate according to the constructed concordance between UCC categories and ELIs (or Item Strata) to get percentile-level household average expenditure $\bar{X}_j^h$ for each 265 ELIs (or 178 items) and the corresponding percentile-level expenditure share $\omega_j^h = \frac{\bar{X}_j^h}{\sum_j \bar{X}_j^h}$.

B.1.2 Constructing Income-Percentile-Specific CPIs

Item-Level Consumer Price Data

To construct the income-percentile-specific consumer price indices (CPI), we need to combine the percentile-level expenditure share data computed above with the micro consumer price data. We obtain the consumer price data from the BLS. Each month, the BLS releases the consumer price index at all levels of aggregation. Each price index has a unique identifier called series id, CUUR0400AA0 for example. The series id can be broken down to: CU–survey abbreviation–current series, U–season code–seasonal unadjusted, R–periodicity code–monthly, 0400–area code–Western Urban and AA0–item code–all items. We use the US city average, all urban consumers, seasonally adjusted item-level monthly price indices to construct the monthly income-percentile-specific CPIs.

Concordance Between Old and New Series

The revised consumer price data were introduced by the BLS in 1998, and the revision included an updated and revised item structure. For example, there were only 7 major groups of goods and services before 1997 and in 1998, a new group *Education and Communication* was created and the new group included components previously included in the *Recreation* and *Housing* groups. Here, we refer to the revised item structure as the new series and to the pre-revised item structure as the old series. Micro consumer price data are provided in the old series before 1997, and in new series since 1997.

To combine the item-level consumer price data from the old series with the expenditure share data, we manually construct a concordance from the new series to the old series at the item level. Note that there are some new series items that are more aggregated than the old ones, and in these cases one item in the new series is concorded to multiple items in old series. To deal with it, we construct a concordance weight by using the expenditure weight taken from the BLS table *Relative Importance of Components in the Consumer Price Index*. One example is as follows. Item *SEFF01 Chicken* in the new series is concorded to *SE0601 Fresh whole chicken* and *SE0602 Fresh/Frozen chicken parts* in the old series. We find that the average expenditure during years 1987 to 1989 on the two items are 0.152% and 0.220% respectively, and thus we assign the concordance weights based on their relative expenditure weights on the two items. The 265 new series items are concorded to 165 old series ones.

item code (new)	item name (new)	item code (old)	item name (old)	exp weight	concordance weight
SEFF01	CHICKEN	SE0601	FRESH WHOLE CHICKEN	0.152	0.409
SEFF01	CHICKEN	SE0602	FRESH/FROZEN CHICKEN PARTS	0.220	0.591

Aggregation Formula

We follow the BLS manual *"Chapter 17. The Consumer Price Index"* in constructing the income-percentile-specific CPI. The formula can be written as follows:

$$PIX_t^h = PIX_v^h \cdot \sum_{j \in J} (\omega_{j,\beta}^h \times \frac{P_{j,t}}{P_{j,v}}),$$

where:
$PIX_t^h =$ consumer price index for household at percentile h at time t
$v =$ pivot year and month, usually December, prior to the month when expenditure weights from reference period (β) are first used in the CPI
$\beta =$ predetermined expenditure reference period
$P_{j,t} =$ price of item j at time t
$\omega_{j,\beta}^h =$ expenditure weights of household at percentile h on item j during the predetermined expenditure reference period β .

The BLS periodically updates its expenditure weight reference period. Historically, it updated approximately every ten years, and since 2002, it adopted a biennial rotation schedule to update the expenditure weight reference period. We follow the BLS expenditure reference period schedule after 2004, and prior to that, we use the 2004 percentile-level expenditure weights to construct the income-percentile-specific CPI. As mentioned in B.1.1, this is due to the availability of the imputed household income before tax. We have computed the pre-2004 aggregate CPI by taking the official expenditure weights from BLS table *Relative Importance of Components in the Consumer Price Index* for the pre-2004 expenditure reference period. And comparing it with the aggregate CPI constructed by using 2004 aggregate weights, we find the two CPI series are almost identical.

Due to the revision of item structure in 1998, we have to construct the income-percentile-specific CPI separately in two periods. We use old series item-level micro price data to compute the income-percentile-specific CPI for the period 1969m1-

1997m12 and new series price data for the post-1998 period. In the year 1997, the BLS released item-level micro prices in both old and new series, which allows us to bridge the two periods by using one of the months in 1997 as the pivot month (based period) for the second period. We used both the old and new series micro price data to construct the aggregate CPIs in 1997 and found that they give us similar results in (log) price terms. We choose 1997m12 as the first pivot month for the construction of the post-1998 income-percentile-specific CPIs.

Table B.2 Reference periods

pivot month (v)	reference period (β)	PIX (t)
1969M1	2004	1969-1997(Old series)
1997M12	2004	1998-2005(New series)
2005M12	2004-2005	2006-2007(New series)
2007M12	2006-2009	2008-2009(New series)
⋮	⋮	⋮
2015M12	2012-2015	2016-2017(New series)

Notes: This table lists the reference periods used to construct the CPI.

B.1.3 Categories with the Largest Expenditure Share Differences

Table B.3 reports the 10 categories with the largest differences in expenditure shares between the top 1% and the middle 20% of households.

Table B.3 Expenditure share differences, frequency of price adjustment, and volatility of price changes

Category	Income percentile		Difference	Regular Price Change	Price Change	St. Dev.
	40-60	100				
Top 10, larger expenditure shares by middle class						
Gasoline (all types)	0.084	0.038	-0.046	87.71	87.74	0.208
Electricity	0.050	0.025	-0.025	38.14	38.14	0.035
Limited service meals and snacks	0.037	0.018	-0.018	6.13	7.00	0.009
Wireless telephone services	0.032	0.014	-0.018	13.00	13.00	0.044
Motor vehicle insurance	0.039	0.021	-0.018	8.16	8.16	0.025
Hospital services	0.037	0.024	-0.013	6.26	6.26	0.014
Cable and satellite television and radio service	0.024	0.012	-0.013	12.35	12.83	0.015
Used cars and trucks	0.028	0.016	-0.012	100.00	100.00	0.052
Prescription drugs	0.022	0.011	-0.011	15.03	15.09	0.015
Cigarettes	0.012	0.001	-0.010	23.17	33.59	0.073
Mean				31.00	32.18	0.049
Median				14.02	14.04	0.030

Table B.3 Continued Expenditure share differences, frequency of price adjustment, and volatility of price changes

Category	Income percentile		Difference	Regular Price Change	Price Change	St. Dev.
	40-60	100				
Top 10, larger expenditure shares by top 1%						
College tuition and fees	0.012	0.051	0.039	5.77	5.77	0.018
Child care and nursery school	0.006	0.030	0.024	6.91	6.91	0.011
Elementary and high school tuition and fees	0.002	0.025	0.023	6.23	6.23	0.013
Watches	0.001	0.021	0.021	3.06	19.83	0.028
Airline fare	0.008	0.028	0.020	59.84	59.84	0.062
Domestic services	0.002	0.019	0.017	4.31	4.31	0.014
Club dues and fees for participant sports and group exercises	0.006	0.022	0.016	8.57	12.56	0.017
Other lodging away from home including hotels and motels	0.007	0.023	0.015	41.73	42.75	0.034
New vehicles	0.048	0.057	0.009	18.89	19.45	0.014
Admissions	0.005	0.013	0.008	8.07	8.39	0.017
Mean				16.34	18.60	0.023
Median				7.49	10.47	0.017

Notes: This table reports the product categories with the largest differences in expenditure shares between the middle (40th-60th percentiles) and the top 1% of the income distribution, the frequency of price changes, and the standard deviation of 12-month log price changes for those products.

B.2 FAVAR evidence

This appendix presents an alternative method to estimate the impulse responses of income-specific CPIs to monetary policy shocks: the Factor-Augmented Vector Autoregression (FAVAR) approach of Bernanke et al. (2005) and Boivin et al. (2009). Let there be a large number of economic series, whose behavior is driven by a vector of common components. This vector includes monetary policy in the form of the Federal Funds rate i_t, and a small number of unobserved common factors $\mathbf{F}_t$. The joint evolution of the Federal Funds rate and the vector of factors, $\mathbf{C}_t$, is characterized by a VAR:

$$\mathbf{C}_t \equiv \begin{bmatrix} \mathbf{F}_t \\ i_t \end{bmatrix},$$

$$\mathbf{C}_t = \Phi(L)\mathbf{C}_{t-1} + \mathbf{v}_t, \tag{B.1}$$

where $\Phi(L)$ is a lag polynomial, and $\mathbf{v}_t$ is an i.i.d. error term.

The vector $\mathbf{F}_t$ is unobservable. What is observed is a large number of economic series $\mathbf{X}_t$. The FAVAR approach assumes that this set of economic series is characterized by a factor model:

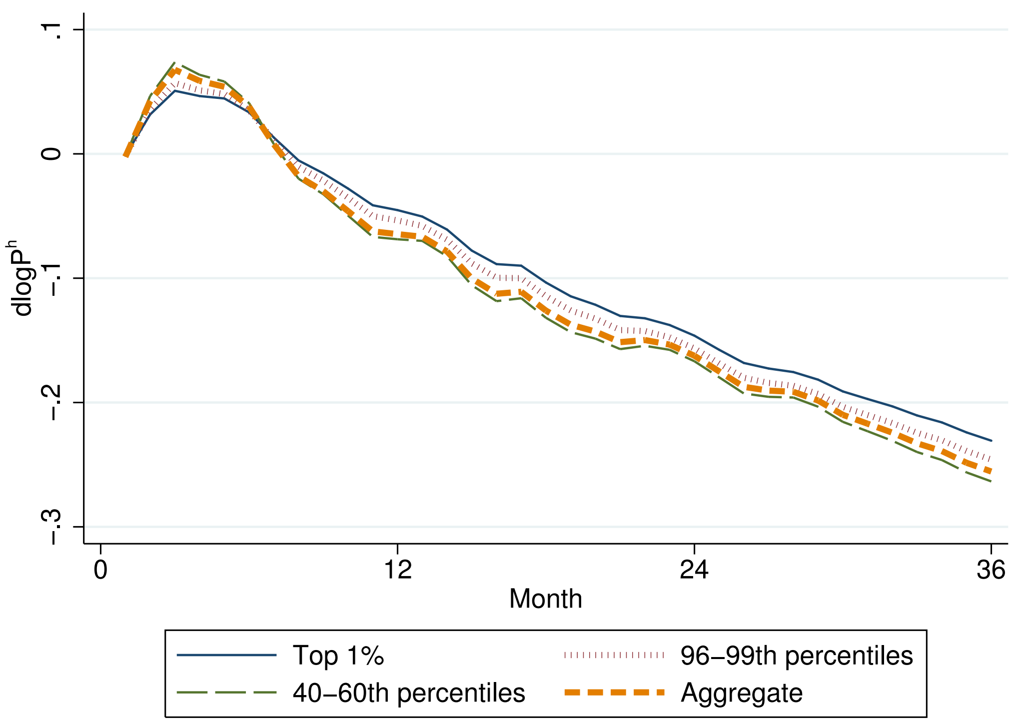
$$\mathbf{X}_t = \mathbf{\Lambda}\mathbf{C}_t + \mathbf{e}_t, \tag{B.2}$$

where Λ is the matrix of factor loadings. This representation provides a great deal of parsimony because in practice $\mathbf{X}_t$ includes hundreds of series, whereas the dimensionality of the vector of common factors $\mathbf{F}_t$ is typically small: in the Boivin et al. (2009) implementation there are 5 common unobserved factors. The significant benefit of estimating model (B.1)-(B.2) is that it yields impulse responses of each of the hundreds of series contained in $\mathbf{X}_t$ to shocks to the elements of $\mathbf{C}_t$, including monetary policy.

In our application of this approach, the vector $\mathbf{X}_t$ includes the 100 income-percentile-specific consumption price indices, as well as the additional variables included by Bernanke et al. (2005) and Boivin et al. (2009), such as sector-level industrial production, employment and earnings, and industry-product-level PPI series. The time frequency is monthly, and the time period is 1978m1-2008m12. Boivin et al. (2009) present a detailed evaluation of the performance of the FAVAR model. Here, we focus on the element new in our paper, namely the impulse responses of income-specific CPIs to monetary policy shocks.

The FAVAR produces 100 of those impulse responses, one for each income percentile. Figure B.1 plots those impulse responses for selected percentiles. The monetary policy shock is a 25-basis-point increase in the Federal Funds rate on impact, thus a contraction. The consumption price indices of the high-income households react substantially less to monetary policy shocks than those for the middle of the income distribution. The difference is economically meaningful. After 12 months, the top-1% households' CPI responds by 34% less, and the 96-99th percentile households by 22% less, than the CPI of the households in the middle of the income distribution (40-60th percentiles). After 24 months, the differences are still 12% and 6%, respectively.

Figure B.1 Income-specific CPI impulse responses to a monetary policy shock

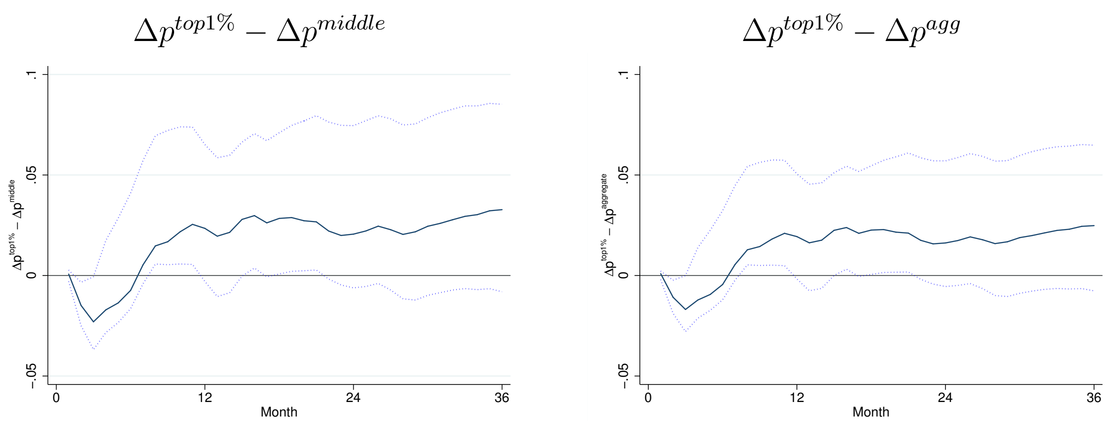


Notes: This figure plots the impulse responses of income-specific price indices to a monetary policy shock, estimated using a FAVAR.

A well-known feature of the VAR impulse responses of prices to monetary shocks is that the confidence intervals are wide, and it is often not possible to reject a zero impact of a monetary shock on aggregate CPI. This is the case in the Boivin et al. (2009) FAVAR model that forms our baseline analysis. However, our main object of interest is not the overall response of prices to a monetary shock, but rather the differential response of the CPIs of different households. Figure B.2 plots the difference in the impulse responses between the CPI of the top 1% and the CPI of the middle

20% of the income distribution (left panel), and the difference between the top 1% and aggregate CPI (right panel). Both panels include the 90% bootstrapped confidence intervals. The difference between impulse responses is significant at the 10% level for most of the lags between 8 and 21 months.[2]

Figure B.2 Differences in inflation changes between income groups



Notes: The left panel plots the difference between the impulse responses of the price index of the top 1% of households and the middle 20% of households to a monetary shock, while the right panel plots the difference between the impulse responses of the price index of the top 1% of households and the aggregate price index, along the 90% bootstrapped confidence intervals.

[2]Note that the impulse is a monetary contractions, and thus the changes in the CPIs are negative after an initial few months. Since the top-income CPIs respond by less in absolute terms, the difference between the top- and middle-income CPIs is positive.

B.3 Substitution Bias

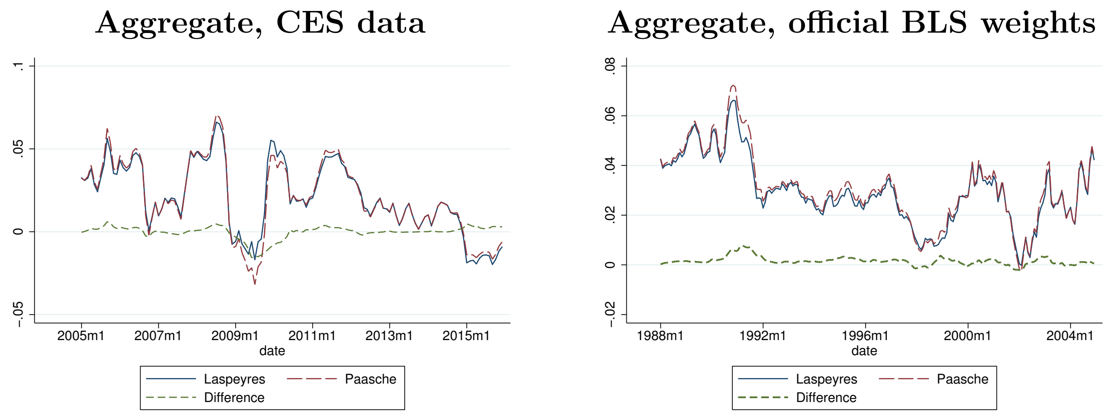
The results in this paper build on the assumption that changes in expenditure shares only have second order effects on inflation. Indeed, the Laspeyres index can be thought of as a first-order approximation to the change in the ideal price index, and thus we rely on the first-order approximation being suitable in this setting. To evaluate this assumption, this Appendix uses year-specific aggregate expenditure shares for each consumption Item from the CES data to construct Laspeyres and Paasche price indices. Since the ideal price index is in-between the Laspeyres and the Paasche, the difference between these two indices provides the upper bound on the bias induced by the first-order approximation.

Figure B.3 below summarizes these differences. It plots 12-month inflation rates for the aggregate CPI computed with Laspeyres and Paasche formulas. From 2004 onwards, we can obtain year-specific aggregate expenditure shares for each consumption Item from the CES data. The CES is the source of expenditure shares data used in the paper. Unfortunately, the product and income definitions in the CES are hard to harmonize prior to 2004. The right panel of Figure B.3 complements the CES data using year-specific aggregate expenditure shares from the BLS between 1987 and 2004 (these expenditure shares are only available from 1987). Both the CES- and the BLS-based measures show little difference between the Laspeyres and the Paasche inflation rates, which confirms that the substitution bias is indeed small in these years. Table B.4 shows that the mean and the standard deviation of the Laspeyres and the Paasche inflation rates are an order of magnitude larger than the mean and standard deviation of the difference between the two measures. The correlation between the Laspeyres and the Paasche inflation rates is 0.99.

Note that the differences between the Laspeyres and the Paasche price indices likely overstate the importance of the substitution bias. Measured expenditure shares may change for reasons other than price changes, such as changes in the composition of households in the expenditure surveys, or changes in tastes across years (see Redding and Weinstein, 2016). In fact, as shown in Table B.4, the standard deviation of the Paasche inflation is larger than of the Laspeyres inflation. Also, there are a number of occasions in which the Paasche inflation is larger than the Laspeyres inflation. This should not be the case if yearly changes in expenditure weights are solely due to substitution towards lower-inflation items.

Figure B.4 uses year-specific expenditure weights by income level computed from the CES to construct Laspeyres and Paasche inflation for the households at the middle

Figure B.3 Aggregate Laspeyres and Paasche CPI inflation



Notes: This figure plots the Laspeyres and Paasche indices, and the difference between the two, for aggregate 12-month CPI inflation. The left panel uses annual aggregate expenditure weights from the CES, the right panel from the BLS.

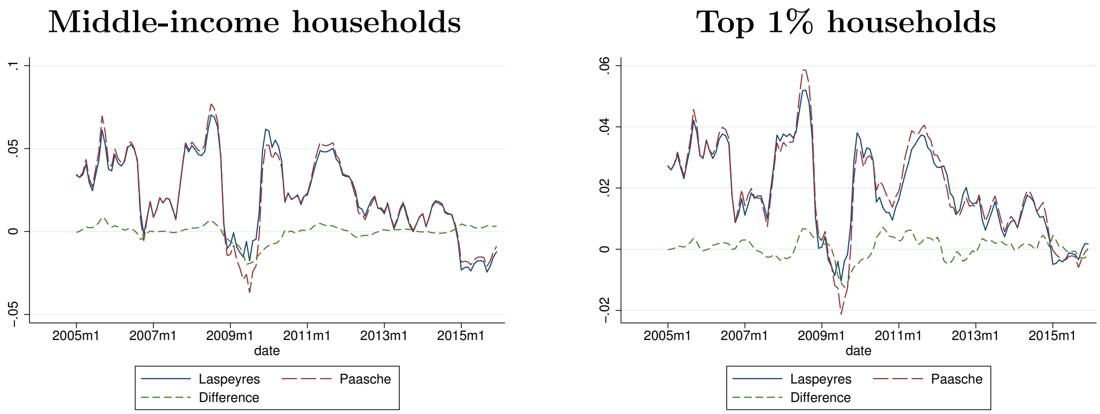
Table B.4 Comparison between the Paasche and Laspeyres price index inflation

	π^L	π^P	$\pi^P - \pi^L$	abs$(\pi^P - \pi^L)$	$Correl(\pi^P, \pi^L)$
1988-2004	2.95%	3.11%	0.16%	0.17%	0.99
	(1.34%)	(1.44%)	(0.17%)	(0.15%)	
2004-2016	2.08%	2.06%	-0.02%	0.25%	0.98
	(2.13%)	(2.26%)	(0.40%)	(0.32%)	

Notes: This table reports the mean and the standard deviation for the Laspeyres price index (π^L), Paasche price index (π^P), the difference between the two ($\pi^P - \pi^L$), and the absolute difference between the two (abs$(\pi^P - \pi^L)$). The last column reports the correlation between the Laspeyres and Paasche inflation rates. The inflation rates are defined as 12-month log changes in the price indices.

and at the top 1% of the income distribution. For the reasons mentioned above, we can only construct these series starting in 2004. Summary statistics for these measures are reported in Table B.5. For each income group, the figure shows that the difference between the Paasche and the Laspeyres inflation is small compared to the overall inflation rates for both groups of households.

Figure B.4 Laspeyres and Paasche CPI inflation by income level



Notes: This figure plots the Laspeyres and Paasche indices, and the difference between the two, for the middle 20% of the households (left panel), and the top 1% of the households (right panel) in the CES.

Table B.5 Comparison between the Paasche and Laspeyres price index inflation, top- and middle-income households

Income	π^L	π^P	$\pi^P - \pi^L$	abs$(\pi^P - \pi^L)$	$Correl(\pi^P, \pi^L)$
Top	1.88%	1.94%	0.05%	0.27%	0.98
	(1.44%)	(1.60%)	(0.35%)	(0.22%)	
Middle	2.16%	2.12%	-0.04%	0.32%	0.98
	(2.36%)	(2.54%)	(0.50%)	(0.38%)	

Notes: This table reports the mean and the standard deviation for the Laspeyres price index (π^L), Paasche price index (π^P), the difference between the two ($\pi^P - \pi^L$), and the absolute difference between the two (abs($\pi^P - \pi^L$)). The last column reports the correlation between the Laspeyres and Paasche inflation rates. The inflation rates are defined as 12-month log changes in the price indices.

Appendix C
Appendices of Chapter III

C.1 Theoretical Appendix

C.1.1 Competitive Equilibrium

The competitive equilibrium of the economy can be represented as the set of solutions to the following system of simultaneous equations:

$$w_j L_{j,k} = \mu_{j,k} \cdot Y_k, \forall j \in J,\ k \in K \tag{C.1}$$

$$\sum_{k \in K} L_{j,k} = \bar{L}_j, \quad \forall j \in J \tag{C.2}$$

$$E = \sum_k \sum_j w_j \cdot L_{j,k} \tag{C.3}$$

$$P_k^{1-\sigma_k} = z_k c_k^{1-\sigma_k} + CMA_k, \quad \forall k \in K \tag{C.4}$$

$$Y_k = c_k^{1-\sigma_k} \left(z_k \frac{e_k \cdot E + \sum_{l \in K} \alpha_{l,k} Y_l}{P_k^{1-\sigma_k}} + FMA_k \right), \quad \forall k \in K \tag{C.5}$$

Here e_k is the fraction of consumer expenditure devoted to industry k, $\mu_{j,k}$ is the fraction of industry k's gross output devoted to purchasing factor input j, and $\alpha_{l,k}$ is the fraction of industry l's gross revenue (Y_l) used to purchase intermediate inputs from sector k. By Shephard's lemma, these shares equal the elasticities of the expenditure or cost functions with respect to the relevant price. Note that these elasticities in principle depend on relative prices, of goods and/or factors. However, homotheticity and (perceived) constant returns imply that they do not depend on total expenditure (E) or industry gross output.

The first set of conditions (C.1) are the industry factor demand equations, which can be summed to generate aggregate factor demand. The second set of conditions (C.2) equates factor demand with fixed factor supply. The third condition equates total factor income and total expenditure, which also ensures (along with the other conditions) that trade balance holds. The fourth set of conditions (C.4) defines the price index, while the fifth set of equations (C.5) defines gross industry revenues as equal to total industry sales.

Notice that the last set of equations can be solved for Y_k as a function of the factor prices and factor allocations (as well as the exogenous market access terms) using matrix algebra. We can then plug this solution into the other equations, and also plug in the definitions of total expenditure and the price indices. We are then left with a set of equations in factor prices and factor allocations. If there is a unique solution for factor allocations given factor prices, i.e. a unique solution $\mathbf{L}$ for the factor demand equations (C.1) given a set of factor prices $\mathbf{w}$, then clearly we can reduce this system to a system of J equating factor demand and factor supply.

In a closed economy, the J equations equating factor supply and demand are homogeneous of degree 1, and hence a normalization is required. In the open economy these equations are not homogeneous of degree 1 in factor prices due to the presence of fixed foreign prices, and no normalization is required.

C.1.2 First Order Welfare Approximation

A general expression for our first order welfare approximation is

$$\begin{aligned} d\ln y = &\sum_{k\in K}\lambda_k^{ex} d\ln FMA_k + \sum_{k\in K}\left(\frac{e_k}{\sigma_k - 1} - \lambda_k^d\right)\theta_k^f d\ln CMA_k \\ &+ d\ln\alpha + \sum_{k\in K}\lambda_k^d d\ln E_k + \sum_{k\in K}\left((1-\sigma_k)(\lambda_k - \lambda_k^d\theta_k^d) - e_k\theta_k^d\right) d\ln c_k, \end{aligned}$$

where λ_k^{ex} (resp. λ_k^d) is the share total sales attributable to industry k's export (resp. domestic) sales, $\lambda_k = \lambda_k^d + \lambda_k^{ex}$, e_k is the consumer expenditure share on industry k, and θ_k^d (resp. θ_k^f) is the share of expenditure on industry k that is sourced domestically (resp. foreign).

Since α, $d\ln c_k$ and $d\ln E_k$ are all ultimately functions of the exogenous variables $d\ln FMA_k$, $d\ln CMA_k$ and $d\ln T_k$, we can substitute in for these variables to derive the expression in the main text.

Planner's Problem

Denote by $q_k^{c,d}$ the quantity of final Home consumption of domestic goods, and by $q_{n,k}^{c,f}$ the quantity of final consumption of foreign goods from country n, and use an i superscript to indicate the corresponding intermediate use. We denote the quantity exported to n by $q^{ex}_{n,k}$, and the production function in each sector by F_k. Define $D_{n,k} \equiv \tau_{n,k}^{1-\sigma_k} E_{n,k}/P_{n,k}^{1-\sigma_k}$.[1]

Using this notation, we can write the planner's problem as

$$\begin{aligned}
\max_{q_k^{c,d}, q_{n,k}^{c,f}, q_k^{i,d}, q_{n,k}^{i,f}, q_{n,k}^{ex}, L_{j,k}} \quad & \ln U(\{q_k^{c,d}\}, \{q_{n,k}^{c,f}\}) \\
s.t. \quad & F_k\left(\{L_{j,k}\}, \{q_k^{i,d}\}, \{q_{n,k}^{i,f}\}\right) = q_k^{c,d} + q_{n,k}^{i,d} + \sum_{n\in N} q_{n,k}^{ex}, \quad \forall k \\
& \sum_k L_{j,k} = \bar{L}_j, \quad \forall j \\
& \sum_k \sum_n p_{n,k}^f \left(q_{n,k}^{c,f} + q_{n,k}^{i,f}\right) = \sum_{k\in K}\sum_{n\in N} (q_{n,k}^{ex})^{\frac{\sigma_k - 1}{\sigma_k}} \cdot D_{n,k}^{\frac{1}{\sigma_k}}.
\end{aligned}$$

We first need to transform this into an expression involving FMA and CMA. Using the first order conditions, it is easy to show that at the optimum for any two export markets n and m

$$\frac{q_{n,k}^{ex}}{q_{m,k}^{ex}} = \frac{D_{n,k}}{D_{m,k}}, \quad \forall m, n \in N,\ k \in K$$

Likewise, from the first order conditions and our CES aggregator for both consumption and intermediate goods, we have

$$\frac{q_{n,k}^{c,f}}{q_{m,k}^{c,f}} = \frac{q_{n,k}^{i,f}}{q_{m,k}^{i,f}} = \left(\frac{p_{n,k}^f}{p_{m,k}^f}\right)^{-\sigma_k}, \quad \forall m, n \in N,\ k \in K$$

This implies that we can define new variables $q_k^{ex} = \sum_{n\in N} q_{n,k}^{ex}$, $q_k^{c,f} = (\sum_{n\in N} (q_{n,k}^{c,f})^{\frac{\sigma_k-1}{\sigma_k}})^{\frac{\sigma_k}{\sigma_k-1}}$ and $q_k^{i,f} = (\sum_{n\in N} (q_{n,k}^{i,f})^{\frac{\sigma_k-1}{\sigma_k}})^{\frac{\sigma_k}{\sigma_k-1}}$ such that the problem above is equivalent to

$$\begin{aligned}
\max_{q_k^{c,d}, q_k^{c,f}, q_k^{i,d}, q_k^{i,f}, q_k^{ex}, L_{j,k}} \quad & \ln U(\{q_k^{c,d}\}, \{q_k^{c,f}\}) \\
s.t. \quad & F_k\left(\{L_{j,k}\}, \{q_k^{i,d}\}, \{q_k^{i,f}\}\right) = q_k^{c,d} + q_{n,k}^{i,d} + q_k^{ex}, \quad \forall k \\
& \sum_k L_{j,k} = \bar{L}_j, \quad \forall j \\
& \sum_k \left(q_k^{c,f} + q_k^{i,f}\right) CMA_k^{\frac{1}{1-\sigma_k}} = \sum_{k\in K} (q_k^{ex})^{\frac{\sigma_k-1}{\sigma_k}} FMA_k^{\frac{1}{\sigma_k}}
\end{aligned}$$

[1]Note that the iceberg assumption implies that the price received by the exporter is

$$p_{n,k}^{ex} = (q_{n,k}^{ex})^{-\frac{1}{\sigma_k}} \cdot D_{n,k}^{\frac{1}{\sigma_k}}.$$

We now derive the formulas for δ_k^{ex} and δ_k^{im} for an efficient economy. A simple application of the Envelope Theorem gives

$$\delta_k^{ex} = \mu \cdot \frac{1}{\sigma_k}, \quad \delta_k^{im} = \mu \cdot \frac{1}{\sigma_k - 1}$$

where μ is the multiplier on the trade balance constraint (and is constant across countries). Our assumption of homotheticity allows us to normalize this constant to equal 1.

Single Factor Economy

We assume upper tier Cobb-Douglas preferences with constant expenditure share e_k. The equilibrium conditions in this case specialize to

$$w\bar{L} = \sum_{k\in K} \left(\frac{w}{T_k}\right)^{1-\sigma_k} \cdot \left(z_k \frac{e_k \cdot w\bar{L}}{z_k \left(\frac{w}{T_k}\right)^{1-\sigma_k} + CMA_k} + FMA_k \right).$$

Taking natural logs of both sides and applying Taylor's theorem with respect to FMA_k and CMA_k, we get

$$d\ln w \approx \sum_{k\in K} \left(\lambda_k^d + (1-\sigma_k)\left(\lambda_k^d \theta_k^f + \lambda_k^{ex}\right)\right) d\ln w + \sum_{k\in K} \lambda_k^{ex} d\ln FMA_k - \sum_{k\in K} \lambda_k^d \theta_k^f d\ln CMA_k$$

The first term captures the effect of changes in wages on domestic costs through both foreign and domestic sales. The second term is the direct effect of changes in export market access. The third term captures the domestic expenditure channel of increases wages. The fourth term captures the effect of changing prices, both domestic and foreign, on nominal income.

Collecting terms and solving for $d\ln w$, we get

$$d\ln w \approx \sum_{k\in K} \frac{\lambda_k^{ex} d\ln FMA_k - \lambda_k^d \theta_k^f d\ln CMA_k}{1 - \sum_{k'\in K}\left(\lambda_{k'}^d + (1-\sigma_{k'})\left(\lambda_{k'}^d \theta_{k'}^f + \lambda_{k'}^{ex}\right)\right)}.$$

To solve for the changes in real income, we need to consider the effect on the overall price index $\mathbb{P} = \prod_{k\in K} P_k^{e_k}$. Using the Cobb-Douglas assumption and the results above, we can write

$$d\ln \mathbb{P} \approx \sum_{k\in K} e_k \left(\theta_k^d d\ln w + \frac{\theta_k^f}{1-\sigma_k} d\ln CMA_k\right)$$

Putting the two results together, we get

$$
\begin{aligned}
d\ln y &\approx d\ln w - \left[\sum_{k\in K}(e_k - \lambda_k^{im})d\ln w + \sum_{k\in K}\lambda_k^{im}\frac{d\ln CMA_k}{1-\sigma_k}\right] \\
&= \lambda^{im} d\ln w - \sum_{k\in K}\lambda_k^{im}\frac{d\ln CMA_k}{1-\sigma_k} \\
&= \lambda^{im}\cdot\sum_{k\in K}\frac{\lambda_k^{ex} d\ln FMA_k - \lambda_k^{d}\theta_k^{f} d\ln CMA_k}{1-\sum_{k'\in K}\left(\lambda_{k'}^{d} + (1-\sigma_{k'})\left(\lambda_{k'}^{d}\theta_{k'}^{f} + \lambda_{k'}^{ex}\right)\right)} - \sum_{k\in K}\lambda_k^{im}\frac{d\ln CMA_k}{1-\sigma_k} \\
&= \kappa\cdot\left[\sum_{k\in K}\lambda_k^{ex} d\ln FMA_k - \frac{\lambda_k^{d}\theta_k^{f}}{\lambda_k^{im}}\lambda_k^{im} d\ln CMA_k\right] - \sum_{k\in K}\lambda_k^{im}\frac{d\ln CMA_k}{1-\sigma_k} \\
&= \kappa\cdot\sum_{k\in K}\lambda_k^{ex} d\ln FMA_k + \sum_{k\in K}\left(\frac{1}{\sigma_k - 1} - \kappa\theta_k^{d}\right)\lambda_k^{im} d\ln CMA_k, \\
\kappa &= \frac{\lambda^{im}}{1-\sum_{k'\in K}\left(\lambda_{k'}^{d} + (1-\sigma_{k'})\left(\lambda_{k'}^{d}\theta_{k'}^{f} + \lambda_{k'}^{ex}\right)\right)}
\end{aligned}
$$

where $\lambda^{im} = \sum_{k\in K}\lambda_k^{im}$.

This expression simplifies to the following when we set the domestic sales share in each industry, θ_k^d, equal to zero:

$$
\begin{aligned}
d\ln y &\approx \kappa\cdot\left[\sum_{k\in K}\lambda_k^{ex} d\ln FMA_k - \frac{1}{1-\sigma_k}\lambda_k^{im} d\ln CMA_k\right], \\
\kappa &= \frac{1}{1-\sum_{k'\in K}(1-\sigma_{k'})\lambda_{k'}^{ex}}.
\end{aligned}
$$

External Economies

We now consider a single factor economy with upper tier Cobb-Douglas preferences (as above), but with external economies of scale as in Kucheryavyy et al. (2018). The cost function in each industry is given by $c_k = \frac{w}{T_k L_k^{\gamma_k}}$. We specialize their model to the case with zero domestic sales in any industry. The equilibrium conditions can be expressed as

$$
w\bar{L} = \sum_{k\in K}\left(\frac{w}{T_k L_k^{\gamma_k}}\right)^{1-\sigma_k}\cdot FMA_k
$$

$$
wL_k = \left(\frac{w}{T_k L_k^{\gamma_k}}\right)^{1-\sigma_k}\cdot FMA_k, \quad \forall k\in K.
$$

We assume that, for all industries, $\gamma_k(\sigma_k - 1) < 1$ to ensure a unique equilibrium that will be interior (and hence exhibit smooth comparative statics). Due to the zero domestic sales assumption, production and consumption are entirely distinct in this economy. Since all consumption is imported, CMA only matters for welfare through its direct impact on the consumption prices, in exactly the same manner as in the case with no spillovers. Hence we focus on production.

Solving the individual factor demand equations for L_k in terms of w and plugging them into the aggregate factor demand = supply equation, we get

$$w\bar{L} = \sum_{k \in K} w^{\frac{(1+\gamma_k)(1-\sigma_k)}{1-\gamma_k(\sigma_k-1)}} \cdot FMA_k^{\frac{1}{1-\gamma_k(\sigma_k-1)}} \cdot T_k^{\frac{\sigma_k-1}{1-\gamma_k(\sigma_k-1)}}$$

Using this expression, it is easy to see that

$$d\ln w \approx \kappa \sum_{k \in K} \left(\frac{1}{1-\gamma_k(\sigma_k-1)} \right) \lambda_k^{ex} d\ln FMA_k$$

where

$$\kappa = \frac{1}{1 - \sum_{k' \in K} \frac{(1+\gamma_{k'})(1-\sigma_{k'})}{1-\gamma_{k'}(\sigma_{k'}-1)} \lambda_{k'}^{ex}}.$$

C.2 Data and Estimation Appendix

C.2.1 Matching the Trade Data to Industries

The international trade data from 1965 to 2015 are from the UN COMTRADE Database, which reports bilateral trade flows at the 4-digit SITC Revision 2 level. To concord the trade data to 1997 NAICS industry classifications, we proceed as follows. First, we assign each 4-digit SITC item to its corresponding 6-digit NAICS industries. For instance, *7511 Typewriters cheque-writing machines* are matched to *333313 Office machinery manufacturing*. Second, for those items that are matched to more than one 6-digit NAICS industries, we check whether it could be assigned to the upper-level 5-digit industry. For example, *8510 Footwear* is matched to *316211 Rubber and plastics footwear manufacturing*, *316212 House slipper manufacturing* and some other 6-digit NAICS industries with the first 5-digits "31612." In this case, we aggregate these 6-digit NAICS industries to the 5-digit one 31621 and concord the 4-digit SITC items to the 5-digit NAICS industry. Third, the same is done for the items that are assigned to more than one 5-digit NAICS industries. We matched them to the corresponding 4-digit NAICS industries.

Overall, the 784 4-digit SITC items are matched to 268 NAICS industries. Among them, 233 industries are in the manufacturing sector, 26 in agriculture, and 9 in mining.

C.2.2 K-means Clustering

Selecting the Number of Clusters with Silhouette Analysis

Rousseeuw (1987) introduces the silhouette plot as a means for clustering evaluation. With this method, each cluster is represented by a silhouette displaying which points lie well within the cluster and which ones are marginal to the cluster. The silhouette plot is based on the silhouette width measure, which compares the similarity (cohesion) of a point to points in its own cluster with the ones in neighboring clusters (separation).

The silhouette width s_i is measured as follows:

i. (Measuring the cohesion) Measuring the average distance between point i and all other points in the same cluster. Denote it as a_i.

ii. (Measuring the separation) Measuring the average distance between i and all points in the nearest cluster. Denote it as b_i.

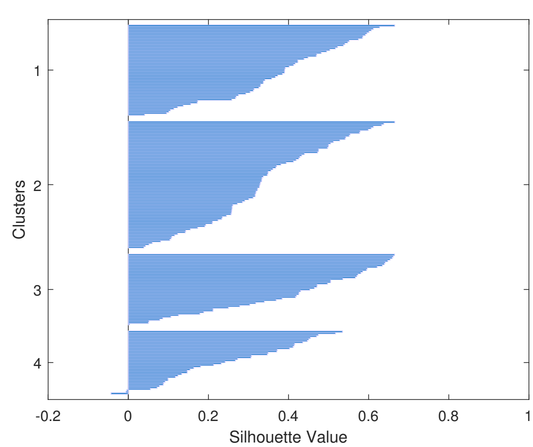
iii. The silhouette width of the observation i is measured as $s_i = \frac{b_i - a_i}{max(a_i, b_i)}$

The silhouette ranges from -1 to 1, where a high value indicates that the point is well assigned to its own cluster and dissimilar to neighboring clusters. A value of 0 indicates that the point is on or very close to the cluster boundary between two neighboring clusters and negative values indicate that those points might have been assigned to the wrong cluster.

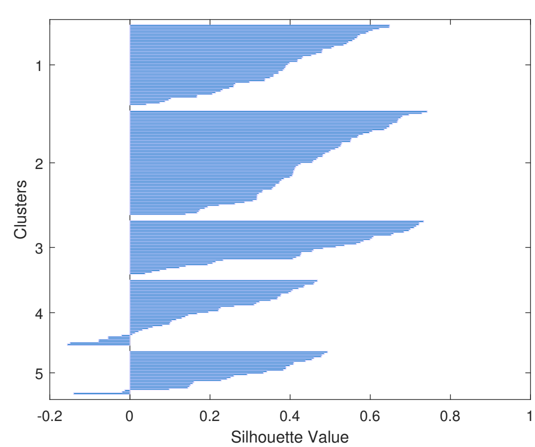
The average silhouette width provides an evaluation of clustering validity, and can be used as way to select an appropriate number of clusters. A high average silhouette width indicates a strong clustering. The average silhouette method computes the average silhouette of observations for different number of clusters G. The optimal number of clusters G is the one that maximizes the average silhouette over a range of possible values for G.

Appendix Figure C.1 plots the silhouette width for industries in each cluster and Appendix Figure C.2 plots the average silhouette over the possible cluster number range. The silhouette analysis suggests that either 4 or 5 are good values for the number of clusters. While the average silhouette value slightly prefers 5 clusters to 4, the silhouette analysis suggests that with 4 clusters fewer industries are near the boundary.

Figure C.1 Silhouette Analysis

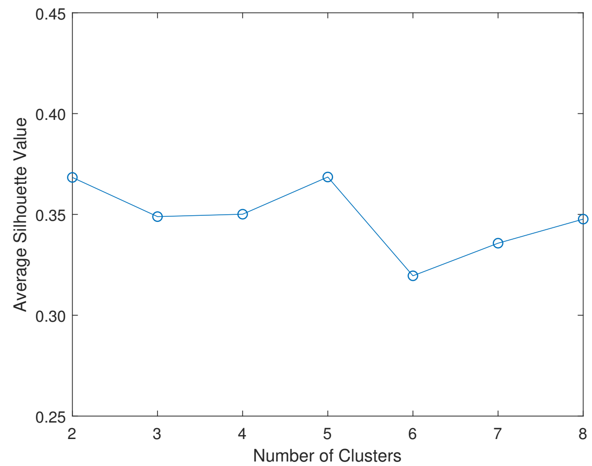


(a) Cluster Number Equals 4



(b) Cluster Number Equals 5

Figure C.2 Average Silhouette Value



Representative Sectors in Each Cluster

The 233 manufacturing sectors are grouped into 4 clusters using the k-means algorithm. Table C.1 lists the 3 most representative sectors in each cluster. The most representative sectors are those closest to the cluster centroid.

Table C.1 The 3 Most Representative Sectors in Each Cluster

Clusters	Label	Representative Sectors	
		Naics	Description
Cluster 1	Raw Materials Processing	324199	All Other Petroleum and Coal Products Manufacturing
		31131	Sugar Manufacturing
		32419	Other Petroleum and Coal Products Manufacturing
Cluster 2	Complex Intermediates	33512	Lighting Fixture Manufacturing
		33531	Electrical Equipment Manufacturing
		339994	Broom, Brush, and Mop Manufacturing
Cluster 3	Capital Goods	333911	Pump and Pumping Equipment Manufacturing
		333994	Industrial Process Furnace and Oven Manufacturing
		333992	Welding and Soldering Equipment Manufacturing
Cluster 4	Consumer Goods	312130	Wineries
		335211	Electric Housewares and Household Fan Manufacturing
		33521	Small Electrical Appliance Manufacturing

K-means Clustering Using a Subset of Characteristic Variables

The average silhouette value of 4 clusters is about 0.35, which indicates that the cluster structure is somewhat weak. However, this could be due to the inclusion of irrelevant sectoral characteristics, which tend to drag down the average silhouette value. We investigate this hypothesis by implementing the algorithm on a subset of important characteristic variables: the investment sales share, intermediates sales shares and contract intensity. These variables are identified as especially important through inspection of the cluster structure as well as more formally using methods developed in Witten and Tibshirani (2010). The 4 clusters based on these three characteristics closely replicate the baseline cluster structure; see Table C.2. The average silhouette value is now about 0.65 (Figure C.3), suggesting a strong cluster structure.

Figure C.3 Average Silhouette Value

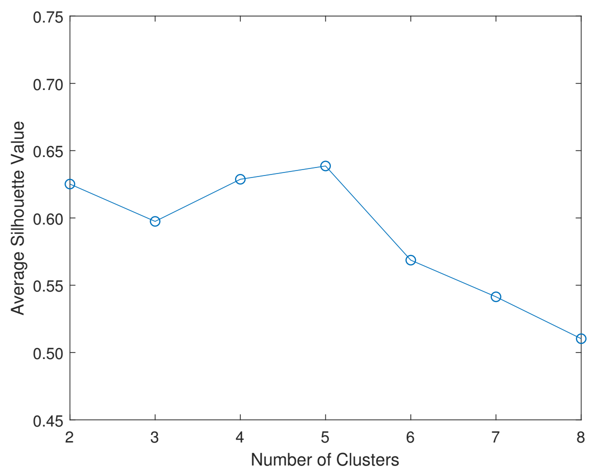


Table C.2 Summary Statistics of Clusters: K-means Clustering Using a Subset of Characteristic Variables

	cluster					
	1	2	3	4	Mean	Std. Dev.
Inv. Share	0.01	0.07	0.56	0.05	0.13	0.22
Int. Using	0.70	0.63	0.65	0.63	0.66	0.16
Int. Sales	0.83	0.78	0.28	0.25	0.57	0.31
Conc. Ratio	0.41	0.30	0.34	0.48	0.40	0.21
Sk. Share	0.30	0.31	0.35	0.33	0.32	0.13
Cap. Int.	0.64	0.55	0.55	0.64	0.61	0.10
Con. Int.	0.29	0.65	0.72	0.57	0.51	0.22
Num of ind.	87	45	42	59		
Trade share	0.38	0.16	0.20	0.19		

C.2.3 Estimation of $FMA_{ik,t}$ and $CMA_{nk,t}$

Equation (3.4) and (3.5) relates external Firm Market Access (FMA) and external Consumer Market Access (CMA) to the gravity equation. The $FMA_{ik,t}$ and $CMA_{nk,t}$ are expressed as follows:

$$
\begin{aligned}
FMA_{ik,t} &= \sum_{n \in N} \frac{E_{n,k}}{P_{n,k}^{1-\sigma_k}} \cdot \tau_{in,k}^{1-\sigma_k}, \\
CMA_{nk,t} &= \sum_{i \in N} c_{i,k}^{1-\sigma_k} \cdot \tau_{in,k}^{1-\sigma_k},
\end{aligned}
$$

where i is exporter and n is importer. The foreign shocks are estimated by using sectoral bilateral trade flow data and a structural gravity equation.

Gravity Regression

Gravity equation (3.2) can be rewritten as

$$
E_{ink,t} \equiv p_{ink,t} \cdot q_{ink,t} = c_{ik,t}^{1-\sigma_k} \cdot \frac{E_{nk,t}}{P_{nk,t}^{1-\sigma_k}} \cdot \tau_{ink,t}^{1-\sigma_k}, \tag{C.6}
$$

where $E_{ink,t}$ denotes country n's total sector k expenditure on goods from country i. We do not observe the domestic trade flows, instead we estimate the share version of this equation *à la* Eaton et al. (2012). Dividing both sides by the total imports of country n, we get

$$
\frac{E_{ink,t}}{\sum_{i \neq n} E_{ink,t}} = c_{ik,t}^{1-\sigma_k} \cdot \frac{E_{nk,t}}{P_{nk,t}^{1-\sigma_k} \cdot \sum_{i \neq n} E_{ink,t}} \cdot \tau_{ink,t}^{1-\sigma_k}.
$$

It can be estimated by regressing bilateral trade flows on exporter and importer fixed effects and bilateral trade distance. The estimating equation is

$$
\ln\left(\frac{E_{ink,t}}{\sum_{i \neq n} E_{ink,t}}\right) = ex_{ik,t} + im_{nk,t} + \zeta_{kt} \ln Distance_{in} + \xi_{kt} Contig_{in} + \epsilon_{ink,t}, \tag{C.7}
$$

where $\frac{E_{ink,t}}{\sum_{i \neq n} E_{ink,t}}$ is the share of total imports from country i to n in sector k at time t, $ex_{ik,t}$ is the exporter fixed effect, $im_{nk,t}$ is the importer fixed effect, ζ_{kt} and ξ_{kt} are the distance and common border coefficients. $Distance_{in}$ measures the geographic distance between country i and n, and $Contig_{in}$ indicates whether country i and n are spatially adjacent.

Importer and exporter fixed effects $im_{nk,t}$ and $ex_{ik,t}$, and the bilateral distance coefficients ζ_{kt} and ξ_{kt} are estimated from the above gravity equation using the Poisson pseudo-maximum likelihood approach of Silva and Tenreyro (2006). The estimation is carried out for each sector and time period separately. We estimate the fixed effects and distance/contiguity coefficients for 127 countries, 268 sectors, and 5 decades spanning 1965-2015.

Shocks to large countries may affect their trading partners' estimated importer and exporter effects. In that case, those estimated fixed effects would not be pure measures of foreign shocks affecting the large country, as they would pick up in part the large country's domestic shocks. To address this potential endogeneity, we carry out the above gravity estimation using the leave-one-out approach. For each country ω, we estimate a set $\{ex_{ik,t}(\omega)\ \ im_{nk,t}(\omega)\ \ \zeta_{kt}(\omega)\ \ \xi_{kt}(\omega)\}$ of country ω-specific exporter and importer fixed effects and distance/contiguity coefficients by dropping country ω from the gravity sample on both the exporter and importer side. In this notation, indexing by ω denotes estimates when country ω is left out of the sample. In practice this does not affect any of our conclusions. The results are very similar if we extract the importer and exporter fixed effects from the simple gravity regression with all countries included. This reflects the fundamental fact that most countries are small in foreign markets.

The fixed effects of log trade flows are identified only up to a sector-time-specific additive constant, and thus we renormalize them by restricting the sum of the importer fixed effects to be zero:

$$\overline{im}_{nk,t}(\omega) = im_{nk,t}(\omega) - \frac{\sum_z im_{zk,t}(\omega)}{N_{kt}(\omega)}$$
$$\overline{ex}_{ik,t}(\omega) = ex_{ik,t}(\omega) + \frac{\sum_z im_{zk,t}(\omega)}{N_{kt}(\omega)},$$

where $N_{kt}(\omega)$ is the total number of countries with positive imports for industry k and time t when ω is left out. In this way, what matters is the share of each country in the total imports across industries, not the total imports of the numéraire country in the fixed effects estimation.

$FMA_{ik,t}$ and $CMA_{nk,t}$

The gravity estimates from the section above can be used to construct $FMA_{ik,t}$ and $CMA_{nk,t}$. The (log) $c_{ik,t}^{1-\sigma_k}$ and $\frac{E_{nk,t}}{P_{nk,t}^{1-\sigma_k} \cdot \sum_{i \neq n} E_{ink,t}}$ are estimated by using the exporter and importer fixed effects respectively. We denote by $\kappa_{ik,t}^{ex}(\omega)$ and $\kappa_{nk,t}^{im}(\omega)$ the estimated

$c_{ik,t}^{1-\sigma_k}$ and $\frac{E_{nk,t}}{P_{nk,t}^{1-\sigma_k} \cdot \sum_{i \neq n} E_{ink,t}}$ when country ω is omitted:

$$\kappa_{ik,t}^{ex}(\omega) = \exp\{\overline{ex}_{ik,t}(\omega)\}$$
$$\kappa_{nk,t}^{im}(\omega) = \exp\{\overline{im}_{nk,t}(\omega)\}.$$

The iceberg bilateral components $\tau_{ink,t}^{1-\sigma_k}$ are estimated by using the bilateral geographic distance and the common border dummy and corresponding distance and common border coefficients. The estimated bilateral component is given by $Distance_{in}^{\zeta_{kt}} \cdot \exp(\xi_{kt} \cdot Contig_{in})$.

The estimated $FMA_{ik,t}$ and $CMA_{nk,t}$ can then be computed as

$$FMA_{ik,t} = \sum_{n \neq i} E_{nk,t}(i) \cdot \kappa_{nk,t}^{im}(i) \cdot Distance_{in}^{\zeta_{kt}(i)} \cdot \exp(\xi_{kt}(i) \cdot Contig_{in})$$
$$CMA_{nk,t} = \sum_{i \neq n} \kappa_{ik,t}^{ex}(n) \cdot Distance_{in}^{\zeta_{kt}(n)} \cdot \exp(\xi_{kt}(n) \cdot Contig_{in}),$$

where $E_{nk,t}(i) \equiv \sum_{i' \neq n,i} E_{i'nk,t}(i)$ is total importer n expenditure when leaving country i out.

C.2.4 The Post-Double-Selection Method

Estimating Equation

The growth estimating equation is specified as follows:

$$d\ln y_{i,t} = \sum_{g \in G} \delta_g^{ex} \cdot [d\ln FMA_{ig,t}] + \sum_{g \in G} \delta_g^{im} \cdot [d\ln CMA_{ig,t}] + \boldsymbol{\eta}\mathbf{w}_{i,t} + \boldsymbol{\phi}\mathbf{s}_{i,t} + D_t + \varepsilon_{i,t},$$

where $d\ln FMA_{ig,t} = \sum_{k \in G} \lambda_{ik,t}^{ex} d\ln FMA_{ik,t}$, $d\ln CMA_{ig,t} = \sum_{k \in G} \lambda_{ik,t}^{ex} d\ln CMA_{ik,t}$ are the log-differenced market access terms aggregated up to the cluster level, and D_t are the time fixed effects.

The vector $\mathbf{w}_{i,t}$ collects the industry-level initial equilibrium variables such as initial import and export shares ($\lambda_{ik,t}^{im}$ and $\lambda_{ik,t}^{ex}$), weighted initial firm and consumer market access ($\lambda_{ik,t}^{ex} \cdot \ln FMA_{ik,t}$ and $\lambda_{ik,t}^{im} \cdot \ln CMA_{ik,t}$), the squares ($(\lambda_{ik,t}^{im})^2$, $(\lambda_{ik,t}^{ex})^2$, $(\lambda_{ik,t}^{ex} \cdot \ln FMA_{ik,t})^2$ and $(\lambda_{ik,t}^{im} \cdot \ln CMA_{ik,t})^2$) and the interactions ($(\lambda_{ik,t}^{ex})^2 \cdot \ln FMA_{ik,t}$ and $(\lambda_{ik,t}^{im})^2 \cdot \ln CMA_{ik,t}$). The vector $\mathbf{s}_{i,t}$ collects the interactions between the initial equilibrium variables and the industry-level foreign shocks, such as $(\lambda_{ik,t}^{ex})^2 \cdot d\ln FMA_{ik,t}$ and $(\lambda_{ik,t}^{im})^2 \cdot d\ln CMA_{ik,t}$.

Since our estimating equation has a large number of controls relative to the sample size, the OLS estimation is infeasible, and dimension reduction is necessary. We estimate the above growth equation by implementing the "post-double-selection" method.

Post-Double-Selection Method

The post-double-selection procedure works in two steps. In the double-selection step, LASSO is applied to select controls variables that are useful for predicting the dependent and independent variables respectively. In the post-selection step, coefficients are estimated via an OLS regression of dependent variables on the independent variables and the selected controls.

First, let's rewrite the estimation equation as follows:

$$d \ln y_{i,t} = \mathbf{d}_{i,t}\boldsymbol{\delta} + \mathbf{x}_{i,t}\boldsymbol{\beta}_y + \mu_{i,t},$$

where $\mathbf{d}_{i,t}$ denotes the vector of treatment variables $d \ln FMA_{ig,t}$ and $d \ln CMA_{ig,t}$, and $\mathbf{x}_{i,t}$ is the vector of control variables.

Applying LASSO directly to our estimation equation above might lead to the omitted-variable bias if the LASSO procedure drops a control variable that is highly correlated with the treatment but the coefficient associated with the control is nonzero. To learn about the relationship between the treatment variables and the controls, let's introduce a reduced-form equation

$$d_{i,t} = \mathbf{x}_{i,t}\boldsymbol{\beta}_d + v_{i,t}$$

for each element $d_{i,t}$ of the vector $\mathbf{d}_{i,t}$.

Substituting the reduced-form $d_{i,t}$ into the growth estimation equation we get

$$\begin{aligned} d \ln y_{i,t} &= \mathbf{x}_{i,t}(\boldsymbol{\beta}_d\boldsymbol{\delta} + \boldsymbol{\beta}_y) + (v_{i,t}\boldsymbol{\delta} + \mu_{i,t}) \\ d_{i,t} &= \mathbf{x}_{i,t}\boldsymbol{\beta}_d + v_{i,t}. \quad \forall d_{i,t} \end{aligned}$$

Both equations are used for variable selection. The first equation is used to select a set of variables that are useful for predicting the dependent variable $d \ln y_{i,t}$ and the second equation is used to select a set of controls that are useful for predicting each of the treatment variables $d_{i,t}$. The reduced form system could be further rewritten as

$$\mathbf{z}_{i,t} = \mathbf{x}_{i,t}\boldsymbol{\beta} + \varepsilon_{i,t}$$

where $\mathbf{z}_{i,t}$ is the vector of dependent variable $d\ln y_{i,t}$ and all treatment variables $d_{i,t}$. A feasible double-selection procedure via LASSO is then defined as follows

$$\min_{\boldsymbol{\beta}} E(\mathbf{z}_{i,t} - \mathbf{x}_{i,t}\boldsymbol{\beta})^2 + \frac{\lambda}{n}||L\boldsymbol{\beta}||_1$$

where $L = diag(l_1, l_2, \ldots, l_p)$ is a diagonal matrix of penalty loadings and λ is the penalty level. The LASSO estimator is used for variable selection by simply selecting the controls with nonzero estimated coefficients.

The double-selection procedure first selects a set of controls that are useful for predicting the independent variable $d\ln y_{i,t}$ and treatment variables $\mathbf{d}_{i,t}$. Then in the post-LASSO step, we estimate δ_g^{ex} and δ_g^{im} by ordinary least squares regression of $d\ln y_{i,t}$ on $\mathbf{d}_{i,t}$ and the union of the variables selected for predicting $d\ln y_{i,t}$ and $\mathbf{d}_{i,t}$.

K-fold Cross Validation

The penalty level λ controls the degree of penalization. Practical choices for λ to prevent overfitting are provided in Belloni et al. (2012, 2014a,b). We follow the online appendix of Belloni et al. (2014a) and choose λ by K-fold cross validation.

The K-fold cross-validation works as follows:

i. Randomly split the data $(y_{i,t}, \mathbf{x}_{i,t}, \mathbf{d}_{i,t})$ into K subsets of equal size, $S_1, S_2, \ldots, S_K$

ii. Set the potential tuning parameter set to be $[\lambda^{RT} - 100 : grid : \lambda^{RT} + 100]$, where $\lambda^{RT} = 2.2\sqrt{n}\Phi(1-\gamma/2p)$ is the rule of thumb tuning parameter suggested in Belloni et al. (2012, 2014b), $\gamma = 0.1/log(p)$, n is the number of observations, p the number of variables, and $grid = 10$.

iii. Given λ, for $k = 1, 2, \ldots, K$:

   (a) *(Training on $(y_{i,t}, \mathbf{x}_{i,t}, \mathbf{d}_{i,t})$, $i \notin S_k$)* Leave the kth subset out, and implement the post-double-selection method with tuning parameter λ on the $K-1$ subsets. Denote the estimated coefficients as $\hat{\boldsymbol{\delta}}^{-k}(\lambda)$ and $\hat{\boldsymbol{\beta}}_y^{-k}(\lambda)$.

   (b) *(Validating on $(y_{i,t}, \mathbf{x}_{i,t}, \mathbf{d}_{i,t})$, $i \in S_k$)* Given $\hat{\boldsymbol{\delta}}^{-k}(\lambda)$ and $\hat{\boldsymbol{\beta}}_y^{-k}(\lambda)$ compute the error in predicting the kth subset,

$$e_k(\lambda) = \sum_{i \in S_k} (d\ln y_{i,t} - \mathbf{d}_{i,t}\hat{\boldsymbol{\delta}}^{-k}(\lambda) - \mathbf{x}_{i,t}\hat{\boldsymbol{\beta}}_y^{-k}(\lambda))^2.$$

iv. This gives the cross-validation error

$$CV(\lambda) = \frac{1}{K} \sum_{1}^{K} e_k(\lambda).$$

v. For each value of the tuning parameter $\lambda \in [\lambda^{RT} - 100, \lambda^{RT} + 100]$, repeat steps 3-4 and choose the tuning parameter that minimizes the $CV(\lambda)$.

Table C.3 Control Variables Selected in the Double-Selection Procedure via LASSO: Baseline Estimation

Controls Included	Controls Selected		
	Baseline	Developed Countries	Developing Countries
$\lambda^{ex}_{ik,t}$	$\lambda^{ex}_{ik104,t}$ $\lambda^{ex}_{ik176,t}$ $\lambda^{ex}_{ik180,t}$	$\lambda^{ex}_{ik64,t}$ $\lambda^{ex}_{ik125,t}$ $\lambda^{ex}_{ik155,t}$ $\lambda^{ex}_{ik165,t}$ $\lambda^{ex}_{ik178,t}$ $\lambda^{ex}_{ik210,t}$ $\lambda^{ex}_{ik241,t}$ $\lambda^{ex}_{ik244,t}$	
$\lambda^{im}_{ik,t}$			
$\lambda^{ex}_{iq,t}$		$\lambda^{ex}_{iq2,t}$	
$\lambda^{im}_{iq,t}$			
$\lambda^{ex}_{ik,t} \cdot \ln FMA_{ik,t}$	$\lambda^{ex}_{ik114,t} \cdot \ln FMA_{ik114,t}$ $\lambda^{ex}_{ik143,t} \cdot \ln FMA_{ik143,t}$ $\lambda^{ex}_{ik179,t} \cdot \ln FMA_{ik179,t}$ $\lambda^{ex}_{ik180,t} \cdot \ln FMA_{ik180,t}$	$\lambda^{ex}_{ik92,t} \cdot \ln FMA_{ik92,t}$ $\lambda^{ex}_{ik180,t} \cdot \ln FMA_{ik180,t}$	
$\lambda^{im}_{ik,t} \cdot \ln CMA_{ik,t}$	$\lambda^{im}_{ik80,t} \cdot \ln CMA_{ik80,t}$ $\lambda^{im}_{ik166,t} \cdot \ln CMA_{ik166,t}$ $\lambda^{im}_{ik172,t} \cdot \ln CMA_{ik172,t}$ $\lambda^{im}_{ik176,t} \cdot \ln CMA_{ik176,t}$ $\lambda^{im}_{ik180,t} \cdot \ln CMA_{ik180,t}$ $\lambda^{im}_{ik186,t} \cdot \ln CMA_{ik186,t}$ $\lambda^{im}_{ik214,t} \cdot \ln CMA_{ik214,t}$ $\lambda^{im}_{ik221,t} \cdot \ln CMA_{ik221,t}$ $\lambda^{im}_{ik222,t} \cdot \ln CMA_{ik222,t}$	$\lambda^{im}_{ik71,t} \cdot \ln CMA_{ik71,t}$ $\lambda^{im}_{ik166,t} \cdot \ln CMA_{ik166,t}$ $\lambda^{im}_{ik180,t} \cdot \ln CMA_{ik180,t}$ $\lambda^{im}_{ik205,t} \cdot \ln CMA_{ik205,t}$ $\lambda^{im}_{ik230,t} \cdot \ln CMA_{ik230,t}$	
$\sum_{k \in q} \lambda^{ex}_{ik,t} \cdot \ln FMA_{ik,t}$			
$\sum_{k \in q} \lambda^{im}_{ik,t} \cdot \ln CMA_{ik,t}$			
Number of Controls Selected	16	16	0
Estimates Figures	Figure 3.1	Figure 3.2	Figure 3.2

Notes: Industries in our sample are relabeled by number from 1 to 281 for coding purpose, i.e. $k = 1, 2, \ldots, 281$. The numbers in the subscripts refers to the corresponding industries.

C.3 Additional Appendix Tables and Figures

Table C.4 Summary Statistics of Clusters: Grouping the Manufacturing Industries to 5 Clusters

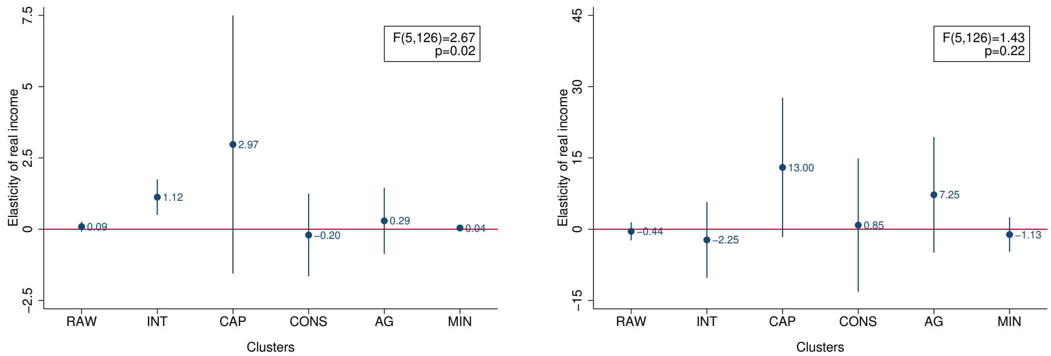
	cluster						
	1	2	3	4	5	Mean	Std. Dev.
Inv. Share	0.00	0.05	0.57	0.03	0.16	0.13	0.22
Int. Using	0.76	0.62	0.67	0.66	0.57	0.66	0.16
Int. Sales	0.85	0.71	0.26	0.31	0.52	0.57	0.31
Conc. Ratio	0.48	0.23	0.35	0.59	0.41	0.40	0.21
Sk. Share	0.33	0.23	0.30	0.32	0.54	0.32	0.13
Cap. Int.	0.69	0.55	0.54	0.69	0.55	0.61	0.10
Con. Int.	0.25	0.52	0.71	0.49	0.74	0.51	0.22
Num of ind.	54	70	36	44	29		
Trade share	0.31	0.20	0.15	0.07	0.20		
Label	Raw Materials	Complex	Capital	Consumer	Skill		
	Processing	Intermediates	Goods	Goods	Intensive		
Abbreviation	RAW	INT	CAP	CONS	SI		

Table C.5 Control Variables Selected in the Double-Selection Procedure via LASSO: Robustness Checks

Controls included	Controls Selected	
Controls Included	Dropping Large Trading Partners	Dropping Contiguous Countries
$\lambda^{ex}_{ik,t}$	$\lambda^{ex}_{ik94,t}$ $\lambda^{ex}_{ik104,t}$ $\lambda^{ex}_{ik114,t}$ $\lambda^{ex}_{ik145,t}$ $\lambda^{ex}_{ik158,t}$ $\lambda^{ex}_{ik176,t}$	$\lambda^{ex}_{ik111,t}$ $\lambda^{ex}_{ik143,t}$ $\lambda^{ex}_{ik176,t}$ $\lambda^{ex}_{ik182,t}$
$\lambda^{im}_{ik,t}$		
$\lambda^{ex}_{iq,t}$		
$\lambda^{im}_{iq,t}$		
$\lambda^{ex}_{ik,t} \cdot \ln FMA_{ik,t}$	$\lambda^{ex}_{ik82,t} \cdot \ln FMA_{ik82,t}$ $\lambda^{ex}_{ik92,t} \cdot \ln FMA_{ik92,t}$ $\lambda^{ex}_{ik94,t} \cdot \ln FMA_{ik94,t}$ $\lambda^{ex}_{ik96,t} \cdot \ln FMA_{ik96,t}$ $\lambda^{ex}_{ik102,t} \cdot \ln FMA_{ik102,t}$ $\lambda^{ex}_{ik143,t} \cdot \ln FMA_{ik143,t}$ $\lambda^{ex}_{ik152,t} \cdot \ln FMA_{ik152,t}$ $\lambda^{ex}_{ik186,t} \cdot \ln FMA_{ik186,t}$ $\lambda^{ex}_{ik190,t} \cdot \ln FMA_{ik190,t}$	$\lambda^{ex}_{ik179,t} \cdot \ln FMA_{ik179,t}$ $\lambda^{ex}_{ik201,t} \cdot \ln FMA_{ik201,t}$ $\lambda^{ex}_{ik203,t} \cdot \ln FMA_{ik203,t}$
$\lambda^{im}_{ik,t} \cdot \ln CMA_{ik,t}$	$\lambda^{im}_{ik127,t} \cdot \ln CMA_{ik127,t}$ $\lambda^{im}_{ik164,t} \cdot \ln CMA_{ik164,t}$ $\lambda^{im}_{ik166,t} \cdot \ln CMA_{ik166,t}$ $\lambda^{im}_{ik176,t} \cdot \ln CMA_{ik176,t}$ $\lambda^{im}_{ik180,t} \cdot \ln CMA_{ik180,t}$ $\lambda^{im}_{ik203,t} \cdot \ln CMA_{ik203,t}$ $\lambda^{im}_{ik217,t} \cdot \ln CMA_{ik217,t}$ $\lambda^{im}_{ik221,t} \cdot \ln CMA_{ik221,t}$ $\lambda^{im}_{ik229,t} \cdot \ln CMA_{ik229,t}$	$\lambda^{im}_{ik166,t} \cdot \ln CMA_{ik166,t}$ $\lambda^{im}_{ik180,t} \cdot \ln CMA_{ik180,t}$ $\lambda^{im}_{ik236,t} \cdot \ln CMA_{ik236,t}$ $\lambda^{im}_{ik237,t} \cdot \ln CMA_{ik237,t}$ $\lambda^{im}_{ik277,t} \cdot \ln CMA_{ik277,t}$
$\sum_{k\in q} \lambda^{ex}_{ik,t} \cdot \ln FMA_{ik,t}$		
$\sum_{k\in q} \lambda^{im}_{ik,t} \cdot \ln CMA_{ik,t}$		
Number of Controls Selected	24	12
Estimates Figures	Figure C.7	Figure C.8

Notes: Industries in our sample are relabeled by number from 1 to 281 for coding purpose, i.e. $k = 1, 2, \ldots, 281$. The numbers in the subscripts refers to the corresponding industries.

Figure C.4 Cluster-Specific Coefficients and Confidence Intervals With a Decreased Tuning Parameter



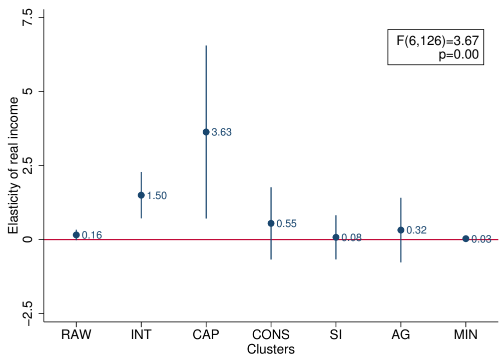
(a) Foreign Demand Shocks

(b) Foreign Supply Shocks

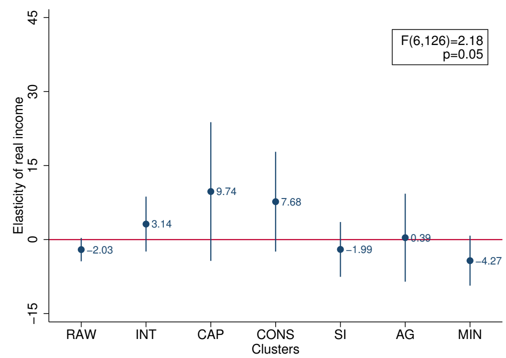
Notes: This figure reports the coefficients in estimating Equation (3.13) with a decreased tuning parameter, for the foreign demand shocks (FMA) (left panels), and foreign supply shocks (CMA) (right panels). The figure displays the post double-LASSO estimates. 38 control variables are selected in the double-selection step. The bars display the 95% confidence bands, that use standard errors clustered by country. The specifications control for initial GDP per capita. The boxes display the results of an F-test for equality of the coefficients in each plot.

Figure C.5 Cluster-Specific Coefficients and Confidence Intervals When Grouping the Manufacturing Industries to 5 Clusters

I. OLS Estimates

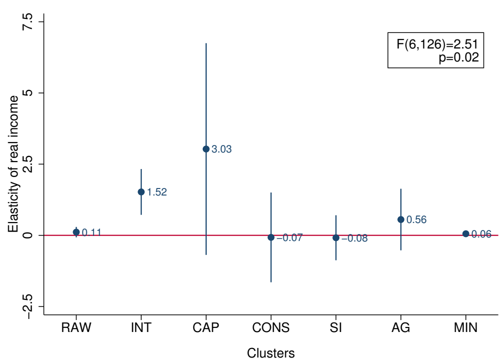


(a) Foreign Demand Shocks

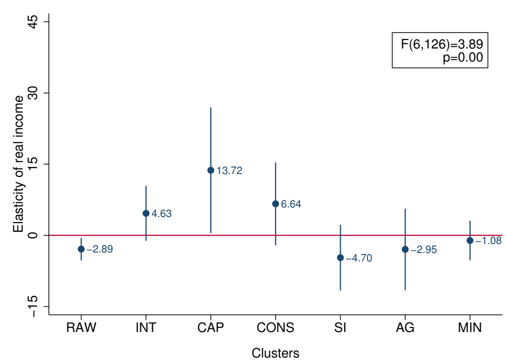


(b) Foreign Supply Shocks

II. LASSO Estimates



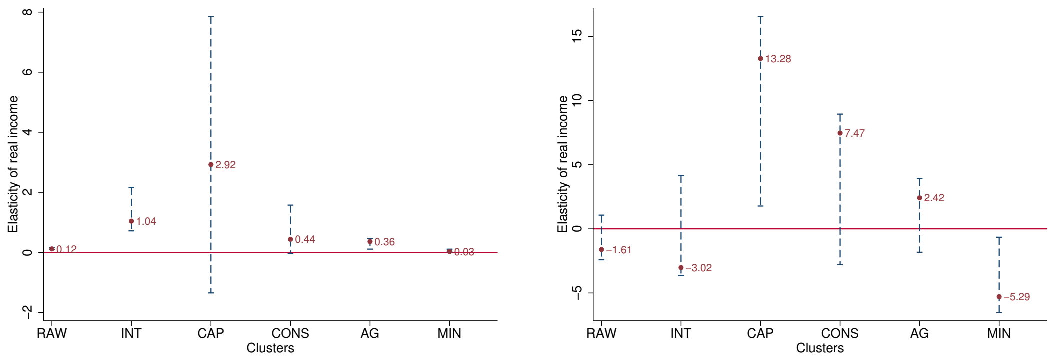
(c) Foreign Demand Shocks



(d) Foreign Supply Shocks

Notes: This figure reports the coefficients in estimating Equation (3.13) when grouping the manufacturing industries to 5 clusters, for the foreign demand shocks (FMA) (left panels), and foreign supply shocks (CMA) (right panels). The top panel displays the baseline OLS estimates. The bottom panel displays the post double-LASSO estimates. 11 control variables are selected in the double-selection step. The bars display the 95% confidence bands, that use standard errors clustered by country. The specifications control for initial GDP per capita. The boxes display the results of an F-test for equality of the coefficients in each plot.

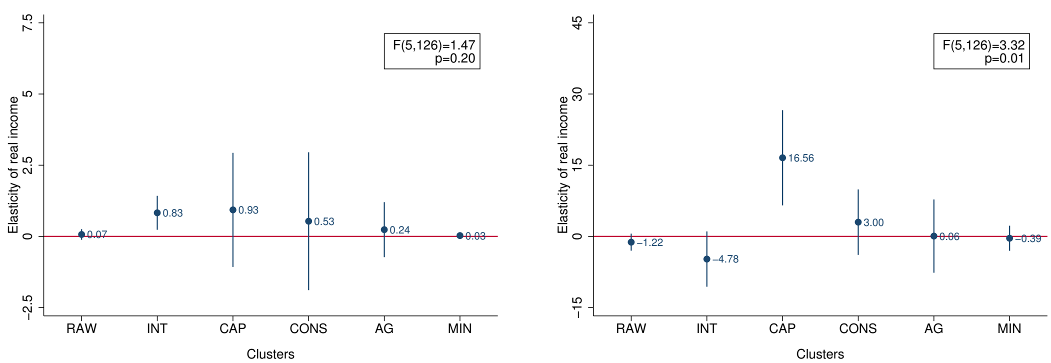
Figure C.6 Cluster Measurement Error Simulation



(a) Foreign Demand Shocks (b) Foreign Supply Shocks

Notes: This figure reports the coefficients in estimating equation (3.13), for the foreign demand shocks (FMA) (left panel), and foreign supply shocks (CMA) (right panel), in the measurement error simulations. The vertical bars report the 95% range of coefficient estimates. The specifications control for initial GDP per capita.

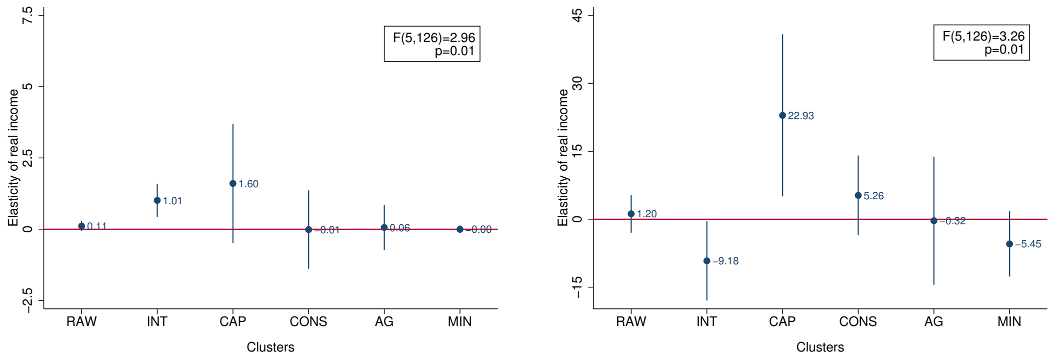
Figure C.7 Dropping Large Trading Partners: Cluster-Specific Coefficients and Confidence Intervals



(a) Foreign Demand Shocks (b) Foreign Supply Shocks

Notes: This figure reports the coefficients in estimating Equation (3.13), for the foreign demand shocks (FMA) (left panels), and foreign supply shocks (CMA) (right panels). The construction of the FMA and CMA terms omit foreign markets for which country i is a large trading partner. The figure displays the post double-LASSO estimates. 24 control variables are selected in the double-selection step. The bars display the 95% confidence bands, that use standard errors clustered by country. The specifications control for initial GDP per capita. The boxes display the results of an F-test for equality of the coefficients in each plot.

Figure C.8 Dropping Contiguous Countries: Cluster-Specific Coefficients and Confidence Intervals



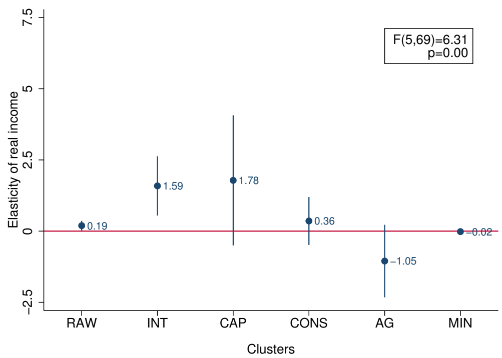
(a) Foreign Demand Shocks

(b) Foreign Supply Shocks

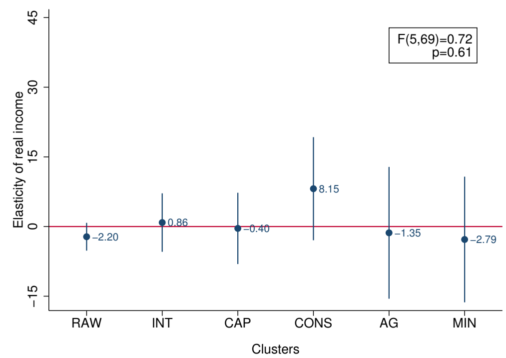
Notes: This figure reports the coefficients in estimating Equation (3.13), for the foreign demand shocks (FMA) (left panel), and foreign supply shocks (CMA) (right panel). The construction of the FMA and CMA terms omit contiguous countries. The figure displays the post double-LASSO estimates. 12 control variables are selected in the double-selection step. The bars display the 95% confidence bands, that use standard errors clustered by country. The specifications control for initial GDP per capita. The boxes display the results of an F-test for equality of the coefficients in each plot.

Figure C.9 Developed vs. Developing Countries: Cluster-Specific Coefficients and Confidence Intervals With a Decreased Tuning Parameter

I. Developed Countries

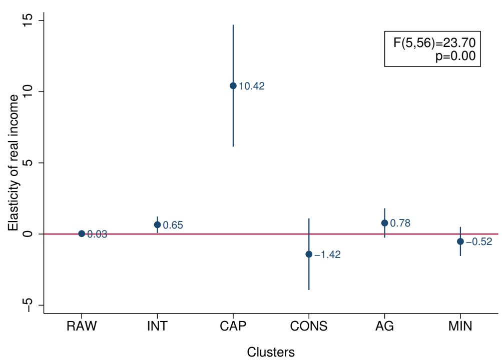


(a) Foreign Demand Shocks

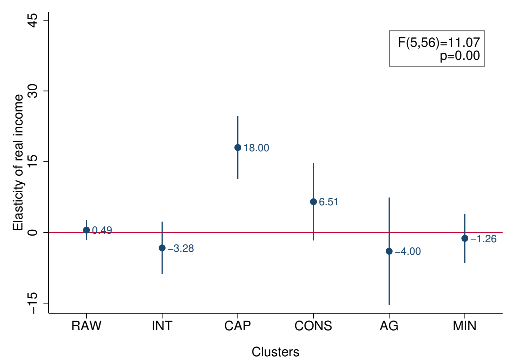


(b) Foreign Supply Shocks

II. Developing Countries



(c) Foreign Demand Shocks

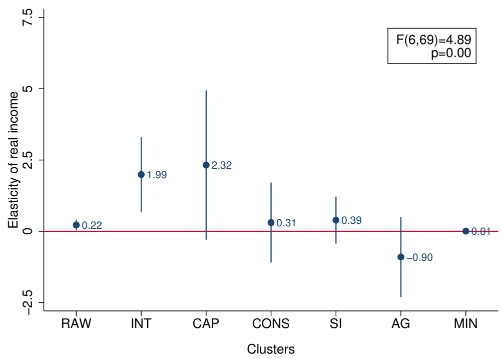


(d) Foreign Supply Shocks

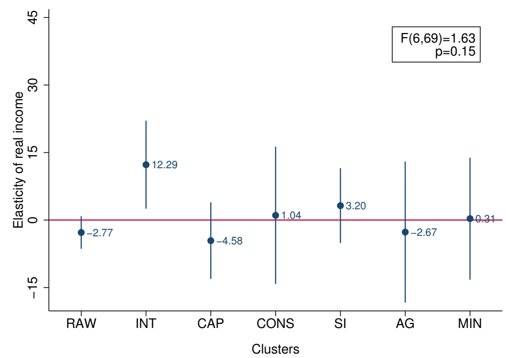
Notes: This figure reports the coefficients in estimating Equation (3.13) with a decreased tuning parameter, for the foreign demand shocks (FMA) (left panels), and foreign supply shocks (CMA) (right panels). The top panel displays the results for the sample of developed countries. 9 control variables are selected in the double-selection step. The bottom panel displays the results for developing countries. 2 control variables are selected in the double-selection step. The bars display the 95% confidence bands, that use standard errors clustered by country. The specifications control for initial GDP per capita. The boxes display the results of an F-test for equality of the coefficients in each plot.

Figure C.10 Developed vs. Developing Countries: Cluster-Specific Coefficients and Confidence Intervals When Grouping the Manufacturing Industries to 5 Clusters

I. Developed Countries

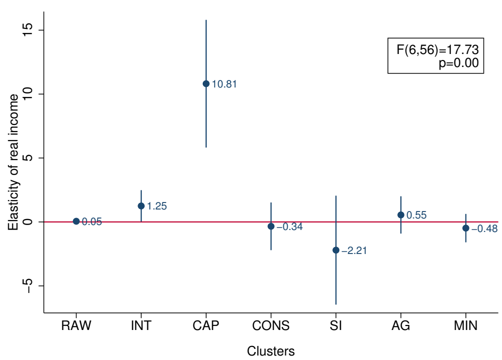


(a) Foreign Demand Shocks

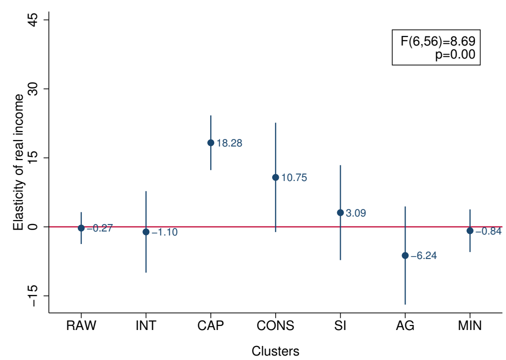


(b) Foreign Supply Shocks

II. Developing Countries



(c) Foreign Demand Shocks

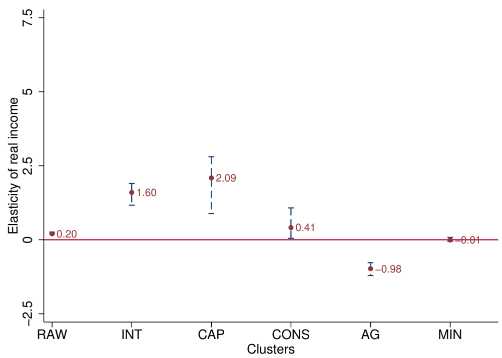


(d) Foreign Supply Shocks

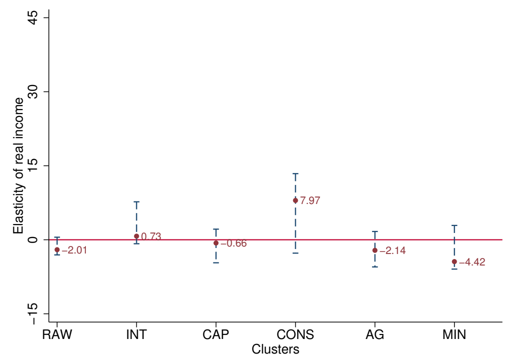
Notes: This figure reports the coefficients in estimating Equation (3.13) when grouping the manufacturing industries to 5 clusters, for the foreign demand shocks (FMA) (left panels), and foreign supply shocks (CMA) (right panels). The top panel displays the results for the sample of developed countries. 14 control variables are selected in the double-selection step. The bottom panel displays the results for developing countries. 6 control variables are selected in the double-selection step. The bars display the 95% confidence bands, that use standard errors clustered by country. The specifications control for initial GDP per capita. The boxes display the results of an F-test for equality of the coefficients in each plot.

Figure C.11 Developed vs. Developing Countries: Cluster Measurement Error Simulation

I. Developed Countries

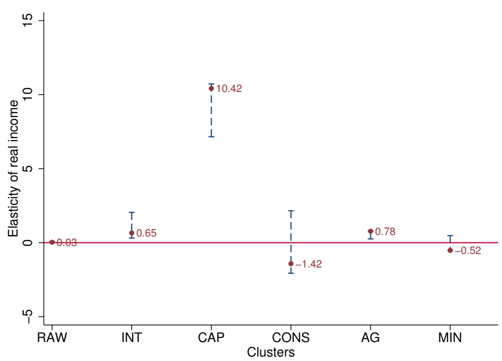


(a) Foreign Demand Shocks

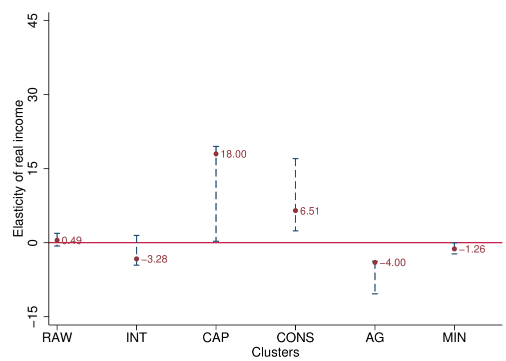


(b) Foreign Supply Shocks

II. Developing Countries



(c) Foreign Demand Shocks

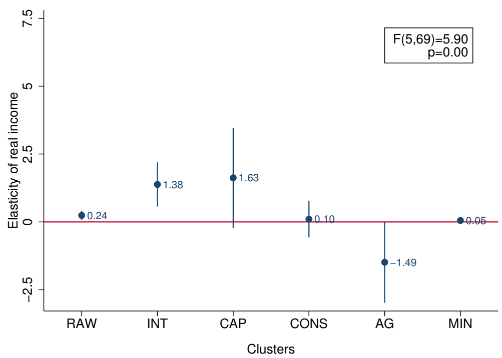


(d) Foreign Supply Shocks

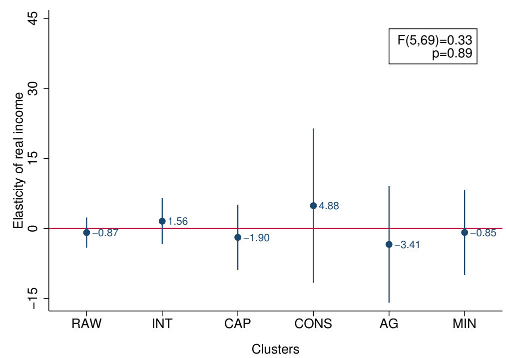
Notes: This figure reports the coefficients in estimating equation (3.13), for the foreign demand shocks (FMA) (left panel), and foreign supply shocks (CMA) (right panel), in the measurement error simulations. The top panel displays the results for the sample of developed countries. The bottom panel displays the results for developing countries. The vertical bars report the 95% range of coefficient estimates. The specifications control for initial GDP per capita.

Figure C.12 Developed vs. Developing Countries: Dropping Large Trading Partners

I. Developed Countries

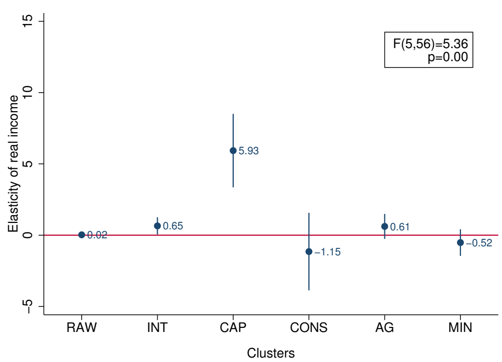


(a) Foreign Demand Shocks

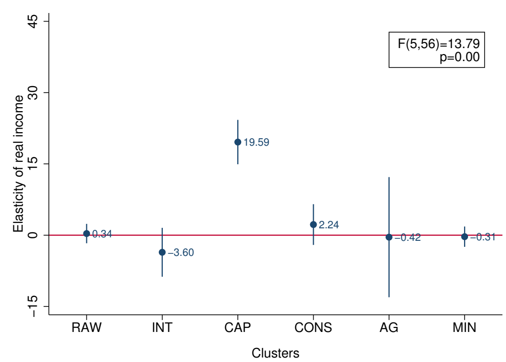


(b) Foreign Supply Shocks

II. Developing Countries



(c) Foreign Demand Shocks

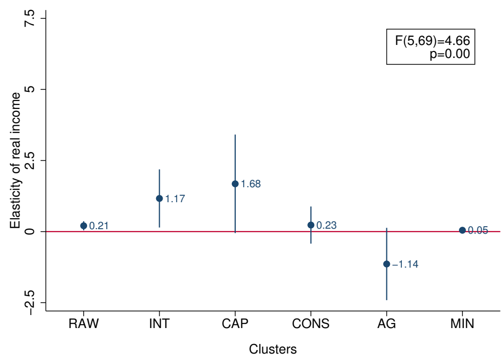


(d) Foreign Supply Shocks

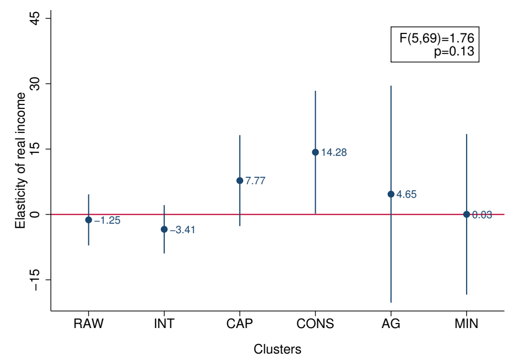
Notes: This figure reports the coefficients in estimating Equation (3.13), for the foreign demand shocks (FMA) (left panels), and foreign supply shocks (CMA) (right panels). The construction of the FMA and CMA terms omit foreign markets for which country i is a large trading partner. The top panel displays the results for the sample of developed countries. 2 control variables are selected in the double-selection step. The bottom panel displays the results for developing countries. 2 control variables are selected in the double-selection step. The bars display the 95% confidence bands, that use standard errors clustered by country. The specifications control for initial GDP per capita. The boxes display the results of an F-test for equality of the coefficients in each plot.

Figure C.13 Developed vs. Developing Countries: Dropping Contiguous Countries

I. Developed Countries

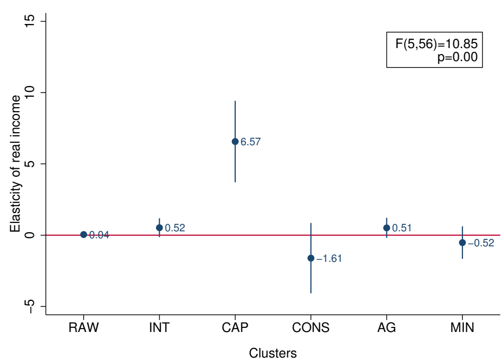


(a) Foreign Demand Shocks

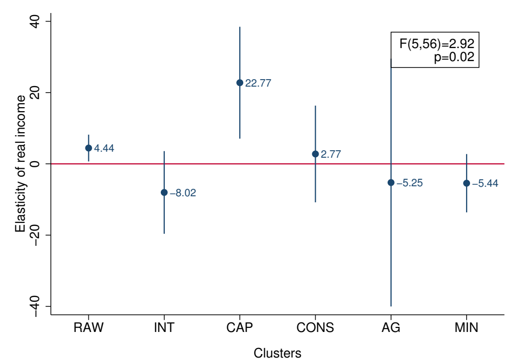


(b) Foreign Supply Shocks

II. Developing Countries



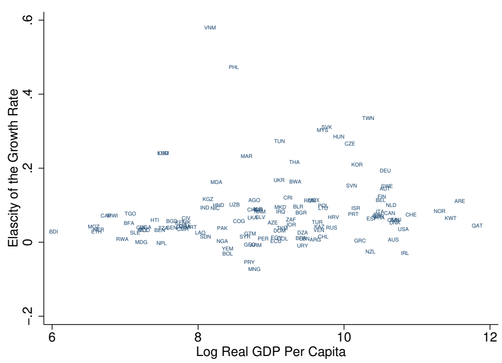
(c) Foreign Demand Shocks



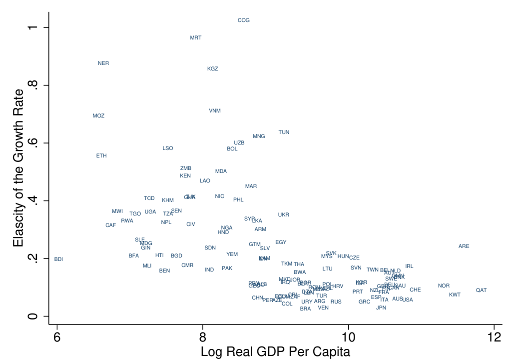
(d) Foreign Supply Shocks

Notes: This figure reports the coefficients in estimating Equation (3.13), for the foreign demand shocks (FMA) (left panels), and foreign supply shocks (CMA) (right panels). The construction of the FMA and CMA terms omit contiguous countries. The top panel displays the results for the sample of developed countries. 3 control variables are selected in the double-selection step. The bottom panel displays the results for developing countries. 1 control variables are selected in the double-selection step. The bars display the 95% confidence bands, that use standard errors clustered by country. The specifications control for initial GDP per capita. The boxes display the results of an F-test for equality of the coefficients in each plot.

Figure C.14 Developed vs. Developing Countries: Elasticity of the Growth Rate



(a) Foreign Demand Shocks



(b) Foreign Supply Shocks

Notes: This figure presents the scatterplot of elasticity of growth rate with respect to the foreign demand shocks (FMA) (left panels), and foreign supply shocks (CMA) (right panels) against real GDP per capita. Elasticity of growth rate is calculated using the developed- and developing-country-specific estimates of coefficients in estimating equation (3.13) and the sectoral export and import shares in 2015.

Bibliography

Bibliography

Adao, Rodrigo, Arnaud Costinot, and Dave Donaldson, "Nonparametric counterfactual predictions in neoclassical models of international trade," *American Economic Review*, 2017, *107* (3), 633–89.

Aghion, Philippe and Xavier Jaravel, "Knowledge spillovers, innovation and growth," *The Economic Journal*, 2015, *125* (583), 533–573.

Aguiar, Angel, Badri Narayanan, and Robert McDougall, "An overview of the GTAP 9 data base," *Journal of Global Economic Analysis*, 2016, *1* (1), 181–208.

Allen, Treb, Costas Arkolakis, and Yuta Takahashi, "Universal gravity," *Forthcoming, Journal of Political Economy*, 2019.

Almås, Ingvild, "International Income Inequality: Measuring PPP Bias by Estimating Engel Curves for Food," *American Economic Review*, 2012, *102* (2), 1093–1117.

Alvaredo, Facundo, Anthony Atkinson, Lucas Chancel, Thomas Piketty, Emmanuel Saez, and Gabriel Zucman, "Distributional National Accounts (DINA) Guidelines : Concepts and Methods used in WID.world," December 2016. WID.world WORKING PAPER SERIES N° 2016/1.

Amiti, Mary and Jozef Konings, "Trade Liberalization, Intermediate Inputs, and Productivity: Evidence from Indonesia," *American Economic Review*, December 2007, *97* (5), 1611–1638.

Argente, David and Munseob Lee, "Cost of Living Inequality during the Great Recession," June 2015. Kilts Booth Marketing Series Paper 1-032.

Arkolakis, Costas, Andrés Rodríguez-Clare, and Jiun-Hua Su, "A multivariate distribution with pareto tails," 2017.

_ , Natalia Ramondo, Andrés Rodríguez-Clare, and Stephen Yeaple, "Innovation and production in the global economy," *American Economic Review*, 2018, *108* (8), 2128–73.

Arthur, David and Sergei Vassilvitskii, "k-means++: The advantages of careful seeding," in "Proceedings of the eighteenth annual ACM-SIAM symposium on Discrete algorithms" Society for Industrial and Applied Mathematics 2007, pp. 1027–1035.

Atkin, David, "Endogenous Skill Acquisition and Export Manufacturing in Mexico," *American Economic Review*, 2016, *106* (8), 2046–2085.

Auclert, Adrien, "Monetary Policy and the Redistribution Channel," May 2017. mimeo, Stanford University.

Audretsch, David B and Maryann P Feldman, "R&D spillovers and the geography of innovation and production," *The American economic review*, 1996, *86* (3), 630–640.

_ and _ , "Knowledge spillovers and the geography of innovation," in "Handbook of regional and urban economics," Vol. 4, Elsevier, 2004, pp. 2713–2739.

Autor, David H, David Dorn, and Gordon H Hanson, "The China syndrome: Local labor market effects of import competition in the United States," *American Economic Review*, 2013, *103* (6), 2121–68.

_ , _ , and _ , "The china shock: Learning from labor-market adjustment to large changes in trade," *Annual Review of Economics*, 2016, *8*, 205–240.

Aw, Bee Yan, Mark J Roberts, and Daniel Yi Xu, "R&D investment, exporting, and productivity dynamics," *American Economic Review*, 2011, *101* (4), 1312–44.

Bartelme, Dominick, "Trade costs and economic geography: evidence from the US," *manuscript*, 2018.

Bartelme, Dominick G, Arnaud Costinot, Dave Donaldson, and Andrés Rodríguez-Clare, "The Textbook Case for Industrial Policy: Theory Meets Data," August 2019. NBER WP 26193.

Bartik, Timothy J, "Who benefits from state and local economic development policies?," *WE Upjohn Institute for Employment Research*, 1991.

Belloni, Alexandre, Daniel Chen, Victor Chernozhukov, and Christian Hansen, "Sparse models and methods for optimal instruments with an application to eminent domain," *Econometrica*, 2012, *80* (6), 2369–2429.

_ , Victor Chernozhukov, and Christian Hansen, "High-dimensional methods and inference on structural and treatment effects," *Journal of Economic Perspectives*, 2014, *28* (2), 29–50.

_ , _ , and _ , "Inference on treatment effects after selection among high-dimensional controls," *The Review of Economic Studies*, 2014, *81* (2), 608–650.

_ , _ , Ivan Fernández-Val, and Christian Hansen, "Program evaluation and causal inference with high-dimensional data," *Econometrica*, 2017, *85* (1), 233–298.

Beraja, Martin, Andreas Fuster, Erik Hurst, and Joseph Vavra, "Regional Heterogeneity and Monetary Policy," August 2017. Forthcoming, *Quarterly Journal of Economics*.

Bernanke, Ben S., Jean Boivin, and Piotr Eliasz, "Measuring the Effects of Monetary Policy: A Factor-Augmented Vector Autoregressive (FAVAR) Approach," *Quarterly Journal of Economics*, 2005, *120* (1), 387–422.

Bilir, L Kamran and Eduardo Morales, "Innovation in the global firm," Technical Report, National Bureau of Economic Research 2016.

Blanchard, Emily J. and William W. Olney, "Globalization and human capital investment: Export composition drives educational attainment," *Journal of International Economics*, 2017, *106*, 165 – 183.

Bloom, Nicholas, Mirko Draca, and John Van Reenen, "Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity," *The Review of Economic Studies*, 2016, *83* (1), 87–117.

Boivin, Jean, Marc P. Giannoni, and Ilian Mihov, "Sticky Prices and Monetary Policy: Evidence from Disaggregated US Data," *American Economic Review*, March 2009, *99* (1), 350–84.

Bøler, Esther Ann, Andreas Moxnes, and Karen Helene Ulltveit-Moe, "R&D, international sourcing, and the joint impact on firm performance," *American Economic Review*, 2015, *105* (12), 3704–39.

Buzard, Kristy, Gerald A Carlino, Robert M Hunt, Jake Carr, Tony E Smith et al., *Localized Knowledge Spillovers: Evidence from the Agglomeration of American R & D Labs and Patent Data*, Federal Reserve Bank of Philadelphia, Research Department, 2015.

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro, "The impact of trade on labor market dynamics," Technical Report, National Bureau of Economic Research 2015.

Calvo, Guillermo A., "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 1983, *12* (3), 383–398.

Capello, Roberta and Camilla Lenzi, "Spatial heterogeneity in knowledge, innovation, and economic growth nexus: conceptual reflections and empirical evidence," *Journal of Regional Science*, 2014, *54* (2), 186–214.

Caragliu, Andrea and Peter Nijkamp, "Space and knowledge spillovers in European regions: the impact of different forms of proximity on spatial knowledge diffusion," *Journal of Economic Geography*, 2015, *16* (3), 749–774.

Carlino, Gerald and William R Kerr, "Agglomeration and innovation," in "Handbook of regional and urban economics," Vol. 5, Elsevier, 2015, pp. 349–404.

Caselli, Francesco and Daniel J. Wilson, "Importing Technology," *Journal of Monetary Economics*, 2004, *51* (1), 1–32.

Christiano, Lawrence J., Mathias Trabandt, and Karl Walentin, "DSGE Models for Monetary Policy Analysis," in Benjamin M. Friedman and Michael Woodford, eds., *Handbook of Monetary Economics*, Vol. 3, Elsevier, 2010, chapter 7, pp. 285–367.

Clayton, Christopher, Xavier Jaravel, and Andreas Schaab, "Heterogeneous Price Rigidities and Monetary Policy," May 2018. Mimeo, Harvard and LSE.

Cohen, Wesley M and Daniel A Levinthal, "Innovation and learning: the two faces of R & D," *The Economic Journal*, 1989, *99* (397), 569–596.

Coibion, Olivier, Yuriy Gorodnichenko, and Gee Hee Hong, "The Cyclicality of Sales, Regular and Effective Prices: Business Cycle and Policy Implications," *American Economic Review*, March 2015, *105* (3), 993–1029.

_ , _ , Lorenz Kueng, and John Silvia, "Innocent Bystanders? Monetary policy and inequality," *Journal of Monetary Economics*, June 2017, *88* (Supplement C), 70 – 89.

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

Cravino, Javier and Andrei A. Levchenko, "The Distributional Consequences of Large Devaluations," *American Economic Review*, November 2017, *107* (11), 3477–3509.

De Long, J Bradford and Lawrence H Summers, "Equipment investment and economic growth," *The Quarterly Journal of Economics*, 1991, *106* (2), 445–502.

Dekle, Robert, Jonathan Eaton, and Samuel Kortum, "Unbalanced trade," *American Economic Review*, 2007, *97* (2), 351–355.

Doepke, Matthias and Martin Schneider, "Inflation and the Redistribution of Nominal Wealth," *Journal of Political Economy*, December 2006, *114* (6), 1069–1097.

Doraszelski, Ulrich and Jordi Jaumandreu, "R&D and productivity: Estimating endogenous productivity," *Review of Economic Studies*, 2013, *80* (4), 1338–1383.

Dorn, David, Gordon H Hanson, Gary Pisano, Pian Shu et al., "Foreign competition and domestic innovation: Evidence from US patents," Technical Report, National Bureau of Economic Research 2016.

Duranton, Gilles and Diego Puga, "Nursery cities: Urban diversity, process innovation, and the life cycle of products," *American Economic Review*, 2001, *91* (5), 1454–1477.

_ and Henry G Overman, "Testing for localization using micro-geographic data," *The Review of Economic Studies*, 2005, *72* (4), 1077–1106.

_ and _ , "Exploring the detailed location patterns of UK manufacturing industries using microgeographic data," *Journal of Regional Science*, 2008, *48* (1), 213–243.

Eaton, Jonathan and Samuel S. Kortum, "Trade in capital goods," *European Economic Review*, 2001, *45* (7), 1195 – 1235.

— and — , "Trade in capital goods," *European Economic Review*, 2001, *45* (7), 1195–1235.

— and — , "Technology, geography, and trade," *Econometrica*, 2002, *70* (5), 1741–1779.

— , — , and Sebastian Sotelo, "International Trade: Linking Micro and Macro," February 2012. NBER WP 17864.

Ellison, Glenn and Edward L Glaeser, "Geographic concentration in US manufacturing industries: a dartboard approach," *Journal of political economy*, 1997, *105* (5), 889–927.

— , — , and William R Kerr, "What causes industry agglomeration? Evidence from coagglomeration patterns," *American Economic Review*, 2010, *100* (3), 1195–1213.

Engel, Ernst, "Die Produktions- und Ernteerträge und der Getreidehandel im preussischen Staate," *Zeitschrift des Königlichen preussischen statischen Bureaus*, 1857, *2*, 249–89.

— , "Das Lebenskosten belgischer Arbeiterfamilien früher und jetzt," *Bulletin de Institut International de Statistique*, 1895, *9*, 1–124.

Faggio, Giulia, Olmo Silva, and William C Strange, "Heterogeneous agglomeration," *Review of Economics and Statistics*, 2017, *99* (1), 80–94.

Feyrer, James, "Trade and Income – Exploiting Time Series in Geography," May 2018. Forthcoming, *Amedican Economic Journal: Applied*.

Frankel, Jeffrey A. and David H. Romer, "Does Trade Cause Growth?," *American Economic Review*, June 1999, *89* (3), 379–399.

Griliches, Zvi et al., "Issues in assessing the contribution of research and development to productivity growth," *Bell Journal of economics*, 1979, *10* (1), 92–116.

Hansen, Lars Peter, "Large sample properties of generalized method of moments estimators," *Econometrica: Journal of the Econometric Society*, 1982, pp. 1029–1054.

Harrison, Ann and Andrés Rodríguez-Clare, "Trade, Foreign Investment, and Industrial Policy for Developing Countries," in Dani Rodrik and Mark Rosenzweig, eds., *Handbook of Development Economics*, New York: North- Holland, 2010, chapter 63.

Hausmann, Ricardo, Jason Hwang, and Dani Rodrik, "What You Export Matters," *Journal of Economic Growth*, 2007, *12*, 1–25.

Head, Keith and John Ries, "Increasing returns versus national product differentiation as an explanation for the pattern of US-Canada trade," *American Economic Review*, 2001, *91* (4), 858–876.

— and Thierry Mayer, "Gravity equations: Workhorse, toolkit, and cookbook," in "Handbook of international economics," Vol. 4, Elsevier, 2014, pp. 131–195.

Hidalgo, Cesar, Bailey Klinger, Albert-Laszlo Barabasi, and Ricardo Hausmann, "The Product Space Conditions the Development of Nations," *Science*, 2007, *371*, 482–487.

Hsieh, Chang-Tai and Ralph Ossa, "A global view of productivity growth in China," *Journal of international Economics*, 2016, *102*, 209–224.

— , Erik Hurst, Charles I Jones, and Peter J Klenow, "The allocation of talent and us economic growth," Technical Report, National Bureau of Economic Research 2013.

Humphreys, Macartan, Jeffrey D. Sachs, and Joseph E. Stiglitz, eds, *Escaping the Resource Curse*, Columbia University Press, 2007.

Jaffe, Adam B, Manuel Trajtenberg, and Rebecca Henderson, "Geographic localization of knowledge spillovers as evidenced by patent citations," *the Quarterly journal of Economics*, 1993, *108* (3), 577–598.

Jaravel, Xavier, "The Unequal Gains from Product Innovations: Evidence from the US Retail Sector," April 2017. mimeo, LSE.

Kaplan, Greg and Sam Schulhofer-Wohl, "Inflation at the household level," *Journal of Monetary Economics*, November 2017, *91*, 19–38.

— , Benjamin Moll, and Giovanni L. Violante, "Monetary Policy According to HANK," *American Economic Review*, March 2018, *108* (3), 697–743.

Kasahara, Hiroyuki and Joel Rodrigue, "Does the use of imported intermediates increase productivity? Plant-level evidence," *Journal of Development Economics*, August 2008, *87* (1), 106–118.

Kim, Seongeun, "Quality, Price Stickiness, and Monetary Policy," April 2018. mimeo, Sejong University.

Krugman, Paul, "Scale economies, product differentiation, and the pattern of trade," *The American Economic Review*, 1980, *70* (5), 950–959.

Kucheryavyy, Konstantin, Gary Lyn, and Andrés Rodríguez-Clare, "Grounded by gravity: A well-behaved trade model with industry-level economies of scale," August 2018. Mimeo, University of Tokyo, Iowa State, and UC Berkeley.

Lagakos, David and Michael E Waugh, "Selection, agriculture, and cross-country productivity differences," *American Economic Review*, 2013, *103* (2), 948–80.

Lederman, Daniel and William F. Maloney, *Does What You Export Matter? In Search of Empirical Guidance for Industrial Policies*, Washington, DC: The World Bank, 2012.

Lee, Jong-Wha, “Capital goods imports and long-run growth,” *Journal of development economics*, 1995, *48* (1), 91–110.

Levinsohn, James and Amil Petrin, “Estimating production functions using inputs to control for unobservables,” *The review of economic studies*, 2003, *70* (2), 317–341.

MacQueen, James et al., “Some methods for classification and analysis of multivariate observations,” in “Proceedings of the fifth Berkeley symposium on mathematical statistics and probability,” Vol. 1(14) 1967, pp. 281–297.

Marshall, Alfred, “Principles of Economics (London, 1920),” *Book VI*, 1920, pp. 618–619.

Melitz, Marc J., “The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity,” *Econometrica*, November 2003, *71* (6), 1695–1725.

Naghavi, Alireza and Gianmarco Ottaviano, “Offshoring and product innovation,” *Economic Theory*, 2009, *38* (3), 517–532.

Nakamura, Emi and Jón Steinsson, “Five Facts about Prices: A Reevaluation of Menu Cost Models,” *Quarterly Journal of Economics*, 2008, *123* (4), 1415–1464.

Nunn, Nathan, “Relationship-specificity, incomplete contracts, and the pattern of trade,” *The Quarterly Journal of Economics*, 2007, *122* (2), 569–600.

_ and Daniel Trefler, “The Structure of Tariffs and Long-Term Growth,” *American Economic Journal: Macroeconomics*, 2010, *2* (4), 158–194.

Olley, G Steven and Ariel Pakes, “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica: Journal of the Econometric Society*, 1996, pp. 1263–1297.

Òscar Jordà, “Estimation and Inference of Impulse Responses by Local Projections,” *American Economic Review*, March 2005, *95* (1), 161–182.

Prebisch, Raúl, “Commercial Policy in the Underdeveloped Countries,” *American Economic Review*, 1959, *49*, 251–273.

Redding, Stephen and Anthony J Venables, “Economic geography and international inequality,” *Journal of international Economics*, 2004, *62* (1), 53–82.

Redding, Stephen J. and David E. Weinstein, “Measuring Aggregate Price Indexes with Demand Shocks: Theory and Evidence for CES Preferences,” 2016. NBER Working Paper 22479.

Rodriguez, Francisco and Dani Rodrik, “Trade Policy and Economic Growth: A Skeptic’s Guide to the Cross-National Evidence,” in “NBER Macroeconomics Annual 2000, Volume 15” NBER Chapters, National Bureau of Economic Research, Inc, November 2001, pp. 261–338.

Rodrik, Dani, "Goodbye Washington consensus, hello Washington confusion? A review of the World Bank's economic growth in the 1990s: learning from a decade of reform," *Journal of Economic literature*, 2006, *44* (4), 973–987.

Romer, Christina D. and David H. Romer, "A New Measure of Monetary Shocks: Derivation and Implications," *American Economic Review*, September 2004, *94* (4), 1055–1084.

Rousseeuw, Peter J, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of computational and applied mathematics*, 1987, *20*, 53–65.

Silva, JMC Santos and Silvana Tenreyro, "The log of gravity," *The Review of Economics and statistics*, 2006, *88* (4), 641–658.

Tecu, Isabel, "The location of industrial innovation: does manufacturing matter?," *US Census Bureau Center for Economic Studies Paper No. CES-WP-13-09*, 2013.

Ventura, Jaume, "A Global View of Economic Growth," in Philippe Aghion and Steven N. Durlauf, eds., *Handbook of Economic Growth*, Vol. 1, Elsevier, 2005, chapter 22, pp. 1419 – 1497.

Whalley, John, *Trade liberalization among major world trading areas*, MIT Press, 1985.

Williamson, Stephen, "Monetary policy and distribution," *Journal of Monetary Economics*, 2008, *55* (6), 1038–1053.

Witten, Daniela M and Robert Tibshirani, "A framework for feature selection in clustering," *Journal of the American Statistical Association*, 2010, *105* (490), 713–726.

Wong, Arlene, "Transmission of Monetary Policy to Consumptionand Population Aging," April 2016. Mimeo, Princeton University.