Distribution of Consumption, Production and Trade within the U.S.

Yilmazkuday, Hakan

2009

Distribution of Consumption, Production and Trade within the U.S.*

Hakan Yilmazkuday†

July 2009

Abstract

This paper attempts to determine the main motivation behind intranational and international trade by introducing a model that considers the distributions of production and consumption *within* the U.S. at the *industry* level. On the consumption side, industry- and state-specific international imports and elasticities of substitution are shown to be systematically connected to consumption agglomeration effects, while on the production side, industry- and state-specific international exports and intermediate input trade are shown to be systematically connected to production agglomeration and specialization effects. Industry structures also play an important role in the determination and magnitude of these effects.

JEL Classification: R12, R13, R32

Key Words: Regional Trade; Intermediate Inputs; The United States

1. Introduction

The current literature in economics is mainly focusing on the international trade and specialization, but much less about domestic (intranational) trade, despite the fact that the latter is orders of magnitude greater than the former. According to the United States (U.S.) trade data, intranational trade volume is more than 3 times international trade volume, on average, between 1993 and 2007.[1] In this context, it would be hard to understand international trade without analyzing first the patterns of intranational trade where there are no additional trade barriers such as tariffs, quotas, cultural differences, language differences, or geography (e.g., Atlantic or Pacific Ocean). If one can figure out the case without these additional constraints (i.e., the intranational trade), it would be easier to analyze the effects of these additional constraints later on (i.e., international trade). In other words, without understanding the patterns of trade in the absence of borders, it is harder to understand them in the presence of borders. In this context, a natural question to ask is "what is the main motivation behind domestic trade?". This paper attempts to answer this question by introducing a model that considers the distributions of *both* production and consumption *within* the U.S. at the *disaggregate* level. Instead of using trade flow data, which do not have sufficient information about the exact distribution (i.e., agglomeration, specialization, etc.) and structure (i.e., technology, marginal costs, etc.) of production and consumption across regions, the consumption, production, and trade (i.e., gross export) implications of a partial equilibrium model are tested using industry-specific production and consumption data at the state level. In particular, four state-level industry data are considered within the U.S.: 1) Food and beverage

*The author would like to thank Eric Bond, Mario Crucini, Kevin Huang, Mark Partridge, John Pepper, Chris Telmer, and Brown Bag Seminar participants at Vanderbilt University for their helpful comments and suggestions. The usual disclaimer applies.

†Department of Economics, Temple University, Philadelphia, PA 19122, USA; Tel: +1-215-204-8880; Fax: +1-215-204-8173; e-mail: hakan.yilmazkuday@temple.edu

[1] Intranational trade data are the sum of all state-level exports (which is equal to the sum of all state-level imports) volume obtained from Commoditiy Flow Survey compiled by the Bureau of Transportation Statistics for the U.S. over the years of 1993, 1997, 2002, 2007. International trade data are the sum of international exports and imports volume obtained from U.S. Census Bureau, Foreign Trade Division, for the same years. The long-run ratio of intranational to international trade volume (which is 3.21) is calculated by taking the average across year-specific ratios which are 4.31 in 1993, 2.03 in 1997, 3.53 in 2002, and 2.97 in 2007.

and tobacco products, 2) Apparel and leather and allied products, 3) Computer and electronic products, and 4) Furniture and related products.[2]

The model consists of individuals and firms in a discrete framework where there are finite number of goods and regions. There are two types of goods, namely traded and non-traded. Each region produces non-traded goods together with a variety of each traded good. Traded goods can be traded up to a transportation cost, and each region may consume varieties of each traded good besides non-traded goods. Production of traded goods is achieved by only labor, while the production of non-traded goods requires traded goods. Thus, traded goods can be used either as a final good or an intermediate input in the model. Individuals in each region have different elasticities of substitution across varieties of each traded good. This, in turn, leads optimization of each monopolistically competitive firm resulting in prices equal to marginal costs with region/good specific mark-ups. According to the model, the main motivation behind trade is found to be the heterogeneity across regions/goods in terms of factor costs, production technologies, transportation technologies, locations, and taste parameters.

Non-traded goods are consumed only locally by definition. So, only the traded goods are modelled in this paper although the existence of non-traded goods, through their intermediate input usage, is considered explicitly. After carefully controlling for intermediate input trade and international trade, the remaining part, the final good trade, is analyzed extensively. In particular, the model is numerically solved using the available data to figure out the region/good specific elasticities of substitution and portions of production that are used as final goods within the country. After that, possible economic connections between international imports, elasticities of substitution, and consumption patterns, as well as connections between international exports, intermediate input trade, and production patterns, are investigated through agglomeration and specialization of the industries at the state level.

Related Literature

The fact that economic geography matters for trade is a well known phenomenon. Nevertheless, modelling the relation between trade and the distribution of economic activity is still in progress. Grossman and Helpman (1995) survey the literature on technology and trade, while Krugman (1980, 1991) provides an introduction to geography and trade via using the economies of scale with transportation costs as the main motivations behind trade. Eaton and Kortum (2002) build a Ricardian model in which the bilateral trade around the world is related to the parameters of geography and technology. Rossi-Hansberg (2005) also builds a spatial Ricardian model, in which, as in Eaton and Kortum (2002), trade is related to the parameters of geography and technology; but this time the technological differences are endogenous and determined by spatial specialization patterns through production externalities. Recently, Alvarez and Lucas (2007) study a variation of the Eaton–Kortum model to investigate the determinants of the cross-country distribution of trade volumes, such as size, tariffs and distance.

The theoretical studies based on gravity equations, such as Anderson (1979), Bergstrand (1985, 1989), among others, also analyze the effects of geography on trade by considering the relation between distance and economic activity across regions. These studies are popular mostly due to their empirical successes.[3] In particular, the first attempt to provide a microeconomic foundation for the gravity models belongs to Anderson (1979). The main motivation behind the gravity model of Anderson (1979) is the assumption that each region is specialized in the production of only one good.[4] Despite its empirical success, as Anderson and van Wincoop (2003) point out, the specialization assumption suppresses finer classifications of goods, and thus makes the model useless in explaining the trade data at the disaggregated level. Another deficiency of the gravity model of Anderson (1979) is the lack of production side. Bergstrand (1985) bridges this gap by introducing a one-factor, one-industry, N-country general equilibrium model in which the production side is considered. In his following study, Bergstrand (1989) extends his earlier gravity model to a two-factor, two-industry, N-country gravity model.[5]

[2] These are the only industries in the U.S. Census Bureau that have both consumption and production data at state level.

[3] Deardorff (1984) reviews the earlier gravity literature. For recent applications, see Wei (1996), Jensen (2000), Rauch (1999), Helpman (1987), Hummels and Levinsohn (1995), and Evenett and Keller (2002).

[4] In appendix of his paper, Anderson (1979) extends his basic model to a model in which multiple goods are produced in each region.

[5] Also see Suga (2007) for a monopolistic-competition model of international trade with external economies of scale, Lopez et al. (2006) for an analysis on home-bias on U.S. iimports of processed food products, and Gallaway et al. (2003) as an empirical study to

Nevertheless, none of the papers mentioned above empirically deal with the patterns of consumption, production, and trade within a country. Recently, Wolf (2000), Hillberry and Hummels (2003), and Yilmazkuday (2009) attempt to bridge this gap by analyzing only the trade patterns by considering trade flow data coming from Commodity Flow Survey (CFS) compiled by the Bureau of Transportation Statistics for the U.S..[6] However, such an analysis would suffer from the lack of actual consumption and production data at the state level, because, as is shown in this paper, agglomeration and specialization of both consumption and production play important roles in the determination of trade patterns, for both final goods and intermediate inputs. Moreover, although CFS is the most available interstate trade data within the U.S., it has deficiencies such as high ratio of missing observations at the industry level. In this context, this paper attempts to employ an alternative measure of trade, total exports, where total exports of a region are broadly defined to include (and distinguish between) intraregional, interregional, or international exports. Moreover, using data of industry level consumption and production obtained from the U.S. Census Bureau at state level, the portions of production that are used as intermediate inputs within the country and/or exported internationally are calculated at the state and industry levels. Figuring out these portions is important as is accepted in the related literature where intermediate inputs have been shown to be playing an important role in trade patterns. In particular, among many others, Hummels, Rapoport, and Yi (1998) document the importance of trade in intermediates; Yi (2003) discusses how trade in intermediates, which implies that a good might cross borders several times during its production, can reconcile the large rise in world trade with relatively modest tariff reductions; Krugman and Venables (1995) provide a model in which, because of trade in intermediates, geography influences the location of industry.

There are also many other regional input-output papers, computable general equilibrium (CGE) papers, or empirical papers based on location quotients (LQs), which have estimated U.S. state-level trade patterns. Compared to these studies, the main contribution of this paper lies in the particular way that the firms/regions are modelled using the monopolistic model, without the need for some of the New Economic Geography (NEG) assumptions. And, most importantly, data support the empirical analyses of the model of this paper with high explanatory powers.

Plan of the Paper

The rest of the paper is organized as follows. Section 2 introduces a regional trade model that explicitly considers intermediate input trade and international trade. Section 3 describes the data used in the empirical analysis. Section 4 employs a state-level empirical analysis to depict the relation between the distribution of consumption and production (through agglomeration or specialization effects) and the portion of production that is used as an intermediate input or exported internationally. Section 5 concludes.

2. The Model

The economy consists of a finite number of goods and a finite number of regions. The model has consumer preferences similar to those continuum-of-goods models that are typical in international trade and open economy macroeconomics studies such as Dornbusch et al. (1977, 1980), Eaton and Kortum (2002), Matsuyama (2000), Erceg et al. (2000), Corsetti and Pesenti (2005), and Gali and Monacelli (2005). In this paper, as in Yilmazkuday (2009), the model adopts this context in a discrete manner, by including heterogeneity across regions/goods in terms of their locations, production technologies, transportation technologies, factor costs, and taste parameters.

The analysis is made for a typical region, r. There are two types of goods, namely traded and non-traded. It is assumed that non-traded goods market is at equilibrium in each region, i.e., consumption of non-traded goods is equal to its production. Since trade implications of the model are of empirical concern, only traded goods are

estimate short-run and long-run industry-level U.S. Armington elasticities.

[6] See Munroe and Hewings (1999) who show that interstate trade is mostly dominated by intra-industry trade. Also see Parr et al. (2002) who suggest that more attention needs to be paid to the mechanisms underlying the manner in which regional economies function and how, over time, greater spatial inter-dependence has become a dominant feature within advanced regional economies of the U.S..

modeled in the analysis, although the existence of non-traded goods are considered explicitly. Each traded good is denoted by $j = 1, ..., J$. Each variety of a traded good is denoted by i which is also the notation for the region producing that variety. In the model, generally speaking, $H_{a,b}(j)$ stands for the variable H, where a is related to the region of consumption, b is related to the variety (and thus, the region of production), and j is related to the good. In this context, $H_{a,b}(j)$ is used for good j of which source location is b and of which destination is a. Needless to say, for presentational purposes, source and destination locations can always be changed, and for instance, $H_{b,a}(j)$ can be used to denote good j of which source location is a and of which destination is b. This notational clarification will be useful especially in the presentation of aggregated variables.

2.1. Individuals

The individual in region r maximizes $\mathbf{U}(\mathbf{C_r})$ where $\mathbf{C_r}$ is a vector of consumption consisting of non-traded goods and traded goods. In region r, consumption of traded good j is given by the following function:[7]

$$C_r(j) = \left(C_r^H(j)\right)^{\gamma_r(j)} \left(C_r^F(j)\right)^{1-\gamma_r(j)} \tag{2.1}$$

where $C_r^H(j)$ is a composite index of good j imported from other regions of the home country, $C_r^F(j)$ is a composite index of good j imported from foreign countries, and, as will be shown below, $\gamma_r(j)$ is the expenditure share of good j that is produced in the home country. $C_r^H(j)$ is further defined as follows:

$$C_r^H(j) \equiv \left(\sum_i (\theta_r(j))^{\frac{1}{\eta_r(j)}} \left(C_{r,i}^H(j)\right)^{\frac{\eta_r(j)-1}{\eta_r(j)}} \right)^{\frac{\eta_r(j)}{\eta_r(j)-1}}$$

where $C_{r,i}^H(j)$ is the variety i of traded good j imported from region i of the home country; $\eta_r(j) > 1$ is the elasticity of substitution across varieties of good j; and finally, $\theta_r(j)$ is a good specific taste parameter.

The optimal allocation of any given expenditure within each variety of traded goods yields the following demand function for goods produced in the home country:

$$C_{r,i}^H(j) = \theta_r(j) \left(\frac{P_{r,i}^H(j)}{P_r^H(j)} \right)^{-\eta_r(j)} C_r^H(j) \tag{2.2}$$

where $P_r^H(j) \equiv \left(\sum_i \theta_r(j) P_{r,i}^H(j)^{1-\eta_r(j)} \right)^{\frac{1}{1-\eta_r(j)}}$ is the price index of traded good j (which is composed of different varieties) that is produced in the home country. When both sides of Equation 2.2 is multiplied by $P_{r,i}^H(j)$ and a summation is taken over i's on both sides, one can obtain the expenditure on traded good j that is produced in the home country as follows:

$$\sum_i P_{r,i}^H(j) C_{r,i}^H(j) = P_r^H(j) C_r^H(j). \tag{2.3}$$

It follows from the optimization (i.e., the cost minimization problem) of Equation 2.1 that:

$$P_r^H(j) C_r^H(j) = \gamma_r(j) P_r(j) C_r(j) \tag{2.4}$$

and

$$P_r^F(j) C_r^F(j) = (1 - \gamma_r(j)) P_r(j) C_r(j) \tag{2.5}$$

where $P_r(j)$ is the price index of traded good j (thus, $P_r(j) C_r(j)$ is the total expenditure on traded good j) in region r, $P_r^H(j)$ is the price index of (thus, $P_r^H(j) C_r^H(j)$ is the expenditure on) traded good j in region r that is produced in the home country, $P_r^F(j)$ is the price index of (thus, $P_r^F(j) C_r^F(j)$ is the expenditure on) traded good j in region r that is produced in foreign countries. Equations 2.4 and 2.5 confirm that $\gamma_r(j)$ is the expenditure share of good j that is produced in the home country.

[7] Type of the utility function and $\mathbf{C_r}$ are irrelevant in the analysis. From now on, unless otherwise stated, goods will refer to traded goods.

2.2. Firms

There are two types of production: (i) traded goods production, (ii) non-traded goods production. While traded goods are produced using sector-specific local labor, non-traded goods are produced using traded goods. In order to have a trackable model, only production in traded goods is introduced, which is sufficient for the empirical analysis of this paper. Nevertheless, the interaction between traded and non-traded goods sectors (i.e., the usage of traded goods as intermediate inputs in the production of non-traded goods) are captured in the market clearing process.

2.2.1. Production of Traded Goods

A typical production firm in region r produces variety r of traded good j using the following production function:

$$Y_r^H(j) = A_r^H(j) L_r^H(j) \tag{2.6}$$

where, $A_r(j)$ represents good and region specific technology, and $L_r(j)$ represents a sector-specific local labor input. The cost minimization problem implies that the marginal cost of producing variety r of good j (in region r) is given by:

$$MC_r^H(j) = \frac{W_r^H(j)}{A_r^H(j)} \tag{2.7}$$

Note that $MC_r(j)$ is good and region specific.

2.2.2. Trade Costs

Anderson and van Wincoop (2004) categorize the trade costs under two names, costs imposed by policy (tariffs, quotas, etc.) and costs imposed by the environment (transportation, wholesale and retail distribution, insurance against various hazards, etc.). Since this paper analyzes trade within a country (i.e., the U.S.), the first category is ignored and the focus is mainly on the second one. Instead of employing a standard "iceberg-melting" trade costs assumption, a unit of traded good j from region r to region i is delivered through a transportation sector. The main difference between an iceberg-melting assumption and having a transportation sector is that additional factors are not used in the production of traded goods; instead, these factors are used in the production of transportation services. By this way, the model has an accurate shipment identity for all traded goods in terms of the market clearing condition. In other words, having a transportation sector is important, because, in the real world (i.e., data), the exporter income is distinguished from the transportation income, which is not the case under the iceberg transport cost assumption. In this context, if there is a trade between regions r and i for good j, trade costs enter the model as follows:[8]

$$\begin{aligned} P_{i,r}^H(j) &= \left(1+\tau_{i,r}^H(j)\right)\left(P_{r,r}^H(j)\right) \\ &= \left(D_{i,r}^H\right)^{\delta(j)}\left(P_{r,r}^H(j)\right) \end{aligned} \tag{2.8}$$

where $P_{r,r}^H(j)$ is the price at the factory gate (i.e., the source); $D_{i,r}^H$ is the distance between regions r and i; and, finally, $\delta(j)$ is good specific elasticity of distance. Here, the expression in the second line in not arbitrary; Yilmazkuday (2008) formally introduces a transportation sector to theoretically connect $(1+\tau_{i,r}(j))$ to $(D_{i,r})^{\delta(j)}$.

2.2.3. Market Clearing

In the model, variety r of good j produced in region r is either (i) consumed domestically in region r (either as a final good or as an intermediate input) or (ii) exported to other regions in the same country (either as a final good or as an intermediate input) or (iii) exported to other countries (either as a final good or as an intermediate input). This condition can be written as:

[8] Note that the existence and volume of trade is determined by Equation 2.2. In particular, it depends on the relative prices of goods imported from different regions as well as the taste parameter, θ.

$$Y_r^H(j) = \sum_i \left(C_{i,r}^H(j) + G_{i,r}^H(j) \right) + \sum_f \left(F_{f,r}^H(j) \right) \tag{2.9}$$

where $C_{i,r}^H(j)$ is consumption of good j as a final good that is produced in region r and consumed in region i (which is in the same country with region r), $G_{i,r}^H(j)$ is consumption of good j as an intermediate input that is produced in region r and consumed in region i (which is in the same country with region r), $F_{f,r}^H(j)$ is consumption of good j either as a final good or an intermediate input that is produced in region r and consumed in foreign country f. In other words, the first term on the right hand side includes intra-regionally consumed good j in region r (when $i = r$) and exported good j to other regions in the same country (when $i \neq r$); the second term on the right hand side consists of international exports of region r. In practice, when $Y_r^H(j)$ represents total shipments rather than total production (the difference of these two is total inventories), Equation 2.9 holds as an accounting identity in equilibrium.

This paper investigates the patterns of intranational trade using state-level U.S. data on total production and final goods consumption at the industry level. In this context, an alternative market clearing condition, this time for final good j consumption within the country, can be written as follows:

$$\alpha_r^H(j)\, Y_r^H(j) = \sum_i C_{i,r}^H(j) \tag{2.10}$$

which is easily obtained by using Equation 2.9 after defining $\alpha_r^H(j)$ as follows:

$$\alpha_r^H(j) = \frac{\sum_i C_{i,r}^H(j)}{\sum_i \left(C_{i,r}^H(j) + G_{i,r}^H(j) \right) + \sum_f \left(F_{f,r}^H(j) \right)} \tag{2.11}$$

where $\alpha_r^H(j)$ is basically the portion of good j production in region r that is consumed as a final good within the home country (i.e., by other regions of the country).

2.2.4. Price Setting for Final Traded Goods

Each firm follows a pricing-to-market strategy in the sense that it sets different prices for final traded goods to be sold in the home country, intermediate traded goods to be sold in the home country, and traded goods (either final or intermediate) to be sold abroad; this paper focuses on the first one. In this context, in region r, the firm producing variety r of final traded good j to be sold in the home country faces the following profit maximization problem:

$$\max_{P_{r,r}(j)} \alpha_r^H(j)\, Y_r^H(j) \left[P_{r,r}^H(j) - MC_r^H(j) \right]$$

subject to Equation 2.10 and the symmetric version of Equation 2.2.[9] The first order condition for this problem implies that:

$$P_{r,r}^H(j) = MC_r^H(j)\, \pi_r^H(j) \tag{2.12}$$

where $\pi_r^H(j)$ represents the gross mark-up:

$$\pi_r^H(j) = \frac{\sum_i \eta_i(j)\, C_{i,r}^H(j)}{\sum_i \left(\eta_i(j) - 1 \right) C_{i,r}^H(j)} \tag{2.13}$$

[9] In an alternative optimization problem, the firm may also maximize its overall profits rather than the profits from final goods to be sold in the home country. In such a case, the optimization result would be the same as long as the firm takes $\alpha_r^H(j)$ as given in its optimization problem.

which is both region and good specific.[10] This is mostly achieved through region specific elasticities of substitution, $\eta_i(j)$'s, rather than a common elasticity of substitution across regions. In a special case in which $\eta_i(j) = \eta(j)$ for all i, the mark-up expression reduces to $\frac{\eta(j)}{\eta(j)-1}$ in all regions. However, data support region specific mark-ups rather than a common mark-up; thus, as in this paper, it is more plausible to use Equation 2.13 in an empirical analysis.[11] Moreover, given the region and good specific mark-ups, together with region and good specific marginal costs (which can be calculated using Equation 2.7), the source prices $P_{r,r}^{H}(j)$ can be obtained through Equation 2.12.

Together with Equation 2.7, Equations 2.12 and 2.13 imply that, for a specific good, the source price differs in each region because of the differences in technology levels, wage rates, sales, and elasticities of substitution in other regions.

2.3. Intraregional and Interregional Trade

According to the model, the nominal value of exports of final traded good j in region r can be written as follows:

$$X_r^H(j) = \alpha_r^H(j) P_{r,r}^H(j) Y_r^H(j) = \sum_i P_{r,r}^H(j) C_{i,r}^H(j)$$

which is basically Equation 2.10 multiplied by the factory gate prices $P_{r,r}(j)$ on both sides. Note that the last expression $\left(\sum_i P_{r,r}^H(j) C_{i,r}^H(j)\right)$ includes both intraregional trade (when $i = r$) and interregional trade (when $i \neq r$). This expression can be rewritten using Equations 2.2 and 2.8 as follows:

$$X_r^H(j) = \alpha_r^H(j) P_{r,r}^H(j) Y_r^H(j) = \sum_i \left(\frac{P_{r,r}^H(j)^{1-\eta_i(j)} \left(\left(D_{i,r}^H \right)^{\delta(j)} \right)^{-\eta_i(j)} P_i^H(j) C_i^H(j)}{\sum_m \left(P_{m,m}^H(j) \left(D_{i,m}^H \right)^{\delta(j)} \right)^{1-\eta_i(j)}} \right) \tag{2.14}$$

Equation 2.14 suggests that the total export of region r for traded final good j depends on the location of each region (due to the trade cost definition in Equation 2.8), the price index of each region (because of the good specific demand functions), the income level of each region (because of the budget constraints) together with elasticities of substitution.

As is well known, the direction of trade could play a crucial role in the distribution of gains from trade under imperfect competition. When Equations 2.7, 2.12 and 2.13 are combined with Equation 2.14, the higher the technology of a region (compared to other regions), the higher are the value of exports. The location of regions are also important through distance measures. To sum up, in order to have a higher volume of exports, a region that is remote from other regions has to compensate its remoteness by having a higher level of technology. This is an important policy implication of the model.

3. Data

Equation 2.14 is empirically tested using state-level industry data within the U.S. These include four 3-digit North American Industrial Classification System (NAICS) industries published by the U.S. Census Bureau for 2002: i) food and beverage and tobacco products, ii) apparel and leather and allied products, iii) computer and electronic products, and iv) furniture and related products.[12] For the rest of the text, food, apparel, electronics, and furniture

[10] Note that the firm has taken aggregate consumption of good j (i.e., $C_i(j)$) and the price index of the traded good j (i.e., $P_i(j)$) as given in all regions (each represented by i) in the optimization problem, because it is relatively small to affect these aggregate variables. In the alternative case in which the firm is not too small, the first order condition becomes less trackable, and no analytical relation can be found between price, marginal cost, and mark-up. In such case, only numerical solution methods can be used to determine the optimal price setting behavior. Although this approach is fine to some extent, when the empirical analysis is considered, a numerical solution is infeasible with the available data.

[11] In particular, according the U.S. Census Bureau production data for 2002, the mark-up values range between 1.17 and 2.83 for food, 1.22 and 7.86 for apparel, 1.17 and 3.10 for electronics, and 1.24 and 2.25 for furniture acros states.

[12] These are the only sectors in the U.S. Census Bureau that have both consumption and production data at state level.

are going to be used respectively, to refer these industries. Because of the data availability, all the states of the United States are included except for Alaska, District of Columbia, and Hawaii. In each industry, the nominal value of *manufacturing* and *retail sales* are used for production in the home country (e.g., $P^H_{r,r}(j)\,Y^H_r(j)$ for region r) and for consumption in the home country (e.g., $P_i(j)\,C_i(j)$ for region i), respectively, in the empirical analysis.[13]

For each industry in each region, to convert consumption in home country (e.g., $P_i(j)\,C_i(j)$ for region i) into consumption that is produced in home country (e.g., $P^H_i(j)\,C^H_i(j)$ for region i), Equation 2.4 is used. Because of the lack of accurate data on international trade at the state level, the consumption shares of good j that is produced in the home country (i.e., $\gamma_r(j)$'s) are available only at the national level (i.e., $\gamma_r(j) = \gamma(j)$ for each industry).[14] In this context, Equation 2.5 can be aggregated across states as follows:

$$\sum_r P^F_r(j)\,C^F_r(j) = (1-\gamma(j))\sum_r P_r(j)\,C_r(j)$$

where $\sum_r P^F_r(j)\,C^F_r(j)$ is the total value of imports of traded good j in the home country (i.e., the U.S.) and $\sum_r P_r(j)\,C_r(j)$ is the total value of consumption of traded good j in the home country (i.e., the U.S.). Using data on retail sales in the home country (i.e., $\sum_r P_r(j)\,C_r(j)$, the total of state-level consumption data obtained from the U.S. Census Bureau for 2002, as introduced above) and the international imports data obtained from the Bureau of Economic Analysis, BEA, (i.e., $\sum_r P^F_r(j)\,C^F_r(j)$, the value of imports given in the national level annual input-output *use* table for 2002), both at the industry level, a value for $(1-\gamma(j))$ is obtained, from which $\gamma(j)$ can be easily calculated for each industry at the national level. In particular, according to data, $\gamma(j) = 0.8955$ for food, $\gamma(j) = 0.2258$ for apparel, $\gamma(j) = 0.1420$ for electronics, and $\gamma(j) = 0.7082$ for furniture. Although data for $\gamma(j)$ values are available only at the national level, the possibility of having state-specific $\gamma_r(j)$ values is discussed in the empirical analysis, and possible implications are further investigated through the model of this paper, below.

In order calculate source prices (i.e., $P^H_{r,r}(j)$'s for all r, j) in Equation 2.14, according to Equations 2.7 and 2.12, industry- and state-specific wage rates, technology levels, and mark-ups are needed. The industry- and state-specific wage rates are obtained from the U.S. Census Bureau data for 2002. The wage rates used are the hourly wage rates of production workers, which are calculated by dividing the the total wage bill of production workers by the average number of hours worked (both data are available at the U.S. Census Bureau). For industry-specific technology levels in each state, the state-level U.S. Census Bureau data for the relevant industries in 2002 are used. In particular, technology level of each industry in each state is proxied by the industry- and state-specific value added (in real terms) per hour of labor. The value added in real terms is calculated by dividing the nominal value added obtained from the U.S. Census Bureau by the cost of living index for each state borrowed from Berry et al. (2003).[15] The industry- and state-specific mark-ups are calculated through dividing total revenue by total costs for each industry in each state using the U.S. Census Bureau data for 2002.

For distance measures, great circle distances between all bilateral states are calculated in statue miles. To calculate the location of each state, the weighted average of latitudes and longitudes of the cities in each state are taken, where the weights are determined according to the production level of those cities. The production level in each city is measured by the real gross domestic product values obtained from BEA for 2002. By using these weights, more relevant spatial locations are obtained for measuring the potential interactions across states. For the distance within each state (i.e., the internal distance), the proxy developed by Wei (1996), which is one-fourth of the distance of a state from the nearest state, is used.

Related to the portion of good j production in region r that is consumed as a final good within the home

[13] A descriptive analysis of these data are available upon request. Such an analysis will also be published at author's personal web page as a supplementary document.

[14] The available international trade data at the state level are recorded according to the location of customs, which do not provide an accurate measure of state-level consumption or production. The reasoning, as also accepted by data collecting agencies, is the fact that the trade of international goods that are recorded at a particular customs in a particular state may be consumed (or might have been produced) in completely another state.

[15] These industry- and state-specific technology levels are available upon request. They will also be presented in the supplementary document of descriptive statistics which will be published in author's personal web page.

country, data for $\alpha_r^H(j)$ are obtained from the annual input-output *use* table of BEA for 2002. However, although these portions are industry specific, they are available only at the national level (i.e., the data cover $\alpha^H(j)$ for all j rather than $\alpha_r^H(j)$ for all r, j). In particular, according to data, $\alpha^H(j) = 0.5969$ for food, $\alpha^H(j) = 0.6267$ for apparel, $\alpha^H(j) = 0.0913$ for electronics, and $\alpha^H(j) = 0.6412$ for furniture. The state-level $\alpha_r^H(j)$'s are numerically solved through the empirical analysis of this paper.

For the inference of empirical results, nominal gross state product (GSP) data for 2002 published by Bureau of Economic Analysis are also used.

Overall, the data set covers each variable in Equation 2.14 except for region and good specific elasticities of substitution across varieties of a good (i.e., $\eta_i(j)$ for all i, j) and good specific elasticities of distance, (i.e., $\delta(j)$ for all j). Instead of assigning specific values for $\eta_i(j)$'s and $\delta(j)$'s, their values are going to be numerically solved according to the model of this paper.

4. Empirical Analysis

Considering the data availability, especially for $\alpha_r^H(j)$ for all r, j, a two-step process is used. For each industry, while the first-step analysis is related to determining the elasticities of substitution across varieties (i.e., the consumption side), the second-step analysis is related to determining the share of output used as a final good within the country (i.e., the production side).

1. First, in order to employ the national-level information for the portion of good j production in region r that is consumed as a final good within the home country (i.e., $\alpha^H(j)$ for all j), Equation 2.14 is aggregated across states to have a national-level expression. In such a case, the only missing parameters are $\eta_i(j)$ for all i, j and $\delta(j)$ for all j; thus, there are totally 48 $\eta_i(j)$'s (one for each state i) and one $\delta(j)$ (totally 49 unknowns) to be determined for each industry j. For each industry, by using 48 state-level mark-ups (i.e., 48 versions of Equation 2.13, one for each state) and one national-level market clearing condition (i.e., the aggregated version of Equation 2.14 across states), these 49 unknowns (i.e., 48 $\eta_i(j)$'s and one $\delta(j)$) can be numerically determined (because there are 49 unknowns and 49 equations). This first-step analysis can be seen as parametrization of the model using the available data and the model. The estimates of $\eta_i(j)$'s are further compared with state-level variables (i.e., industry-specific consumption agglomeration and specialization) to check for possible relations. This is a very similar exercise with Trefler's (1995) experiment in which he solves for the technology levels of the countries and then looks for a possible correlation between technology levels and wage rates. In sum, given the model and $\alpha^H(j)$ for all j, the first-step analysis not only provides estimates of $\eta_i(j)$ for all i, j and $\delta(j)$ for all j, but also depicts the empirical implications of these estimates which are important in understanding consumption and trade patterns of individuals at the state level.

2. Second, using the results of the first-step analysis (i.e., numerically solved 48 $\eta_i(j)$'s and one $\delta(j)$ for each state), the model is tested at the state level using Equation 2.14. Because of the lack of state-level data, $\alpha_r^H(j)$ (for all r, j) are numerically solved using Equation 2.14 (where, for each industry, there are 48 unknown $\alpha_r^H(j)$'s and 48 versions of Equation 2.14, one for each state). The calculated $\alpha_r^H(j)$ (for all r, j) are then compared with state-level variables (i.e., industry-specific production agglomeration and specialization) to check for possible relations. In sum, given the model, numerically solved $\eta_i(j)$'s, and $\delta(j)$ for each state and industry, the second-step analysis not only provides estimates of $\alpha_r^H(j)$ for all r, j, but also depicts the empirical implications these estimates which are important in understanding the production and trade patterns of firms at the state level.

4.1. First-Step Analysis

For the first-step analysis, for each industry j, the aggregation of Equation 2.14 across 48 states results in:

$$\alpha^{H}(j)\sum_{r}P_{r,r}^{H}(j)\,Y_{r}^{H}(j)=\sum_{r}\sum_{i}\left(\frac{P_{r,r}^{H}(j)^{1-\eta_{i}(j)}\left(\left(D_{i,r}^{H}\right)^{\delta(j)}\right)^{-\eta_{i}(j)}P_{i}^{H}(j)\,C_{i}^{H}(j)}{\sum_{m}\left(P_{m,m}^{H}(j)\left(D_{i,m}^{H}\right)^{\delta(j)}\right)^{1-\eta_{i}(j)}}\right) \tag{4.1}$$

where $\alpha^{H}(j)$ is the national-level portion of industry j production that is consumed as a final good within the home country (i.e., the U.S.) that satisfies:

$$\alpha^{H}(j)=\frac{\sum_{r}\alpha_{r}^{H}(j)\,P_{r,r}^{H}(j)\,Y_{r}^{H}(j)}{\sum_{r}P_{r,r}^{H}(j)\,Y_{r}^{H}(j)}=\frac{\sum_{r}\sum_{i}\alpha_{r}^{H}(j)\,P_{r,r}^{H}(j)\,C_{i,r}^{H}(j)}{\sum_{r}P_{r,r}^{H}(j)\,Y_{r}^{H}(j)} \tag{4.2}$$

Using Equation 2.2, Equation 2.13 can be written as follows:

$$\pi_{r}^{H}(j)=\frac{\sum_{i}\eta_{i}(j)\left(\frac{\left(P_{r,r}^{H}(j)\left(D_{i,r}^{H}\right)^{\delta(j)}\right)^{-\eta_{i}(j)}P_{i}^{H}(j)C_{i}^{H}(j)}{\sum_{m}\left(P_{m,m}^{H}(j)\left(D_{i,m}^{H}\right)^{\delta(j)}\right)^{1-\eta_{i}(j)}}\right)}{\sum_{i}\left(\eta_{i}(j)-1\right)\left(\frac{\left(P_{r,r}^{H}(j)\left(D_{i,r}^{H}\right)^{\delta(j)}\right)^{-\eta_{i}(j)}P_{i}^{H}(j)C_{i}^{H}(j)}{\sum_{m}\left(P_{m,m}^{H}(j)\left(D_{i,m}^{H}\right)^{\delta(j)}\right)^{1-\eta_{i}(j)}}\right)} \tag{4.3}$$

So, in this first-step analysis, for each industry j, there are 49 equations (i.e., one from Equation 4.1 and 48 from 4.3, one for each state) in order to determine 49 unknowns (i.e., one $\delta(j)$, and 48 $\eta_i(j)$'s, one for each state i). Using a numerical solution method, nonlinear least squares (NLS), these 49 unknowns for each industry j are exactly identified via available data.

Since empirically tested expressions of this paper are nonlinear, the selection of the starting values in determining the NLS parameters (i.e., one $\delta(j)$, and 48 $\eta_i(j)$'s, one for each state i) are important. Recall that in a special case in which $\eta_i(j)=\eta(j)$ for all i, the mark-up expression reduces to $\frac{\eta(j)}{\eta(j)-1}$ in all regions. Using the average mark-up (where average is taken across states), $\eta(j)$ can be calculated for each industry and used as a starting value for the estimation of all $\eta_i(j)$'s. In particular, the starting value of $\eta_i(j)$'s for food are set to 2.659, for apparel to 2.070, for electronics to 2.636, and for furniture to 2.911. The starting value of $\delta(j)$ is set to a very small number (i.e., $\delta(j)=0.0001$) to allow for a large set of possibilities.

Empirical Results of the First-Step Analysis

A summary of the results is given in Table 1. Although the median elasticity of substitution measures $\eta_i(j)$ are somehow close to each other, the elasticity of distance measures are significantly different from each other across industries. The elasticity of distance takes its highest value for electronics and the lowest for apparel. High transportation cost values for furniture and low values for food and especially apparel are reasonable when their physical structure (especially, their weight and volume) is considered. However, high transportation cost values for electronics is surprising. A possible explanation, for sure, comes from the details of the electronics industry in the data set. According to the data, electronics industry includes the manufacturing of low-weight and/or low-volume equipment (such as compact disks, audio tapes, etc.) as well as high-weight and/or high-volume equipment (such as satellite antennas, coin-operated jukeboxes, loudspeakers, magnetic resonance imaging (MRI) medical diagnostic equipment, ultrasonic medical equipment, radar systems, automatic teller machines (ATMs), etc.). When these details are considered together with the fragile structure of the electronics goods, high transportation cost values for electronics also become reasonable.

In order to explain why these results make sense, one needs to be clear regarding exactly what is meant by distance (i.e., trade) costs. The distance (i.e., transport cost) elasticities $\delta(j)$ are not iceberg specifications, but rather based on a more orthodox view of distance costs as a distance costs mark-up on the price at the factory gate (Equation 2.8). However, nor are they simply transport costs (i.e. movement costs) mark-ups based on weight and distance only, for which the values would not differ significantly across all four industries. Rather they are distance costs, which determine the delivered price. On this point, if one has a distance costs mark-up, then the overall costs of distance transportation are reflected in the final delivered price. But these costs also include all of the inventory holding logistics and shipping costs, all of which are related to shipment frequency which itself is determined by both the product value-weight ratio and the product density as well as the transport (movement) costs. Taking this broad view of distance costs, products which are either very low density per ton (furniture) or very high value per ton (electronics) have very high distance costs. This has been explained analytically by McCann (1998, 2001) and demonstrated empirically in the case of the electronics sector by McCann and Fingleton (1996). On the other hand, products with relatively low density per ton (apparel) or low value per ton (food) exhibit low distance costs, which is consistent with the argument above.

The highest median elasticity of substitution belongs to furniture, while the lowest one belongs to apparel. The complete vector of $\eta_i(j)$'s that include state specific measures for each industry are given in Figures 1-4. As is evident, elasticity of substitution ranges between 2.63 and 2.68 for food, 2.03 and 2.13 for apparel, 2.55 and 3.15 for electronics, and 2.86 and 2.95 for furniture. Although these regional differences between $\eta_i(j)$'s are not substantial, they are sufficient to explain mark-up differences (each given by Equation 2.13 or Equation 4.3), where mark-up values range between 1.17 and 2.83 for food, 1.22 and 7.86 for apparel, 1.17 and 3.10 for electronics, and 1.24 and 2.25 for furniture, across states. Another observation is that the state-level $\eta_i(j)$'s do not seem to follow a geographical pattern. Nevertheless, in order to analyze for possible economic connections, they are compared to other state specific variables in Table 2. While the state-level industry-specific consumption is to capture the agglomeration effects, the state-level industry-specific consumption clustering is to capture the specialization effects. The terms of agglomeration and specialization are generally used for production patterns (as will be discussed in more details below), but, here, following Hoch (1972), Nordhaus and Tobin (1972), Kelley (1977), Fogarty and Garofalo (1980), Power (1981), Tabuchi (2000), and Glaeser et al. (2000), they are used for consumption patterns. As is evident, state- and industry-level elasticities of substitution are highly correlated with state-level consumption of the same industry (i.e., the agglomeration effect). In particular, $\eta_i(j)$'s for food and furniture are positively correlated with $P_i^H(j)\,C_i^H(j)$'s, while they are negatively correlated for apparel and electronics. One possible explanation lies under the structures of these industries: while food and furniture can be seen as more homogenous (which is also supported by the median elasticities of substitution given in Table 1), apparel and electronics may be seen as more heterogenous. More specifically, the elasticity of substitution increases with consumption for food and furniture, because higher consumption of a more homogenous product brings higher elasticities due to the high search and long-distance commuting costs, via agglomeration (see Hoch, 1972, Nordhaus and Tobin, 1972). Similarly, the elasticity of substitution decreases with consumption for apparel and electronics, because higher consumption of a heterogenous product brings lower elasticities due to information spillover among individuals related to the differences (i.e., selectivity) across varieties, via agglomeration. As is also evident, state- and industry-level elasticities of substitution are weakly correlated with state-level consumption clustering of the same industry (i.e., the specialization effect). Nevertheless, the structure of the industries (in terms of their homogeneity) may still work as a possible explanation, except for food.

Analyzing the correlation coefficients does not depict the exact relation between elasticity of substitution and agglomeration and specialization. Also, it is hard to make a comparison across industries with only correlation coefficients. Moreover, agglomeration effects can be correlated to specialization effects which would make the isolation of their individual effects harder. In order to figure out these details, a regression analysis is employed including these variables. The results are given in Table 3. As is evident, the agglomeration effects of consumption are significant for all industries, while the specialization effects of consumption are significant only for apparel and electronics. In particular, across states of the U.S., 1 percent increase in industry-specific consumption corresponds to 0.005 percent rise in the elasticity of substitution for food, 0.009 percent fall for apparel, 0.065 percent fall in electronics, and 0.007 percent rise for furniture. The high coefficient estimate for electronics (especially, relative to apparel) seems to reflect the high degree of information spillover in the context discussed above. The significant specialization effects for apparel and electronics support the idea that individuals relatively consuming more apparel and electronics benefit more from information spillover, while there is no such evidence for food and furniture. When both agglomeration and specialization effects are considered, they both become insignificant, mostly due to multicollinearity between agglomeration and specialization effects. Finally, the explanatory powers of the regressions are high, which support the analysis.

In sum, the elasticities of substitution (i.e., $\eta_i(j)$'s) change across states, and these changes can be systematically explained by the structure of the products together with the distribution of industry-specific consumption within the country. Since the elasticity of substitution is a key parameter that is used by policy makers to derive quantitative results (because the effects of a policy change are evaluated by the conversion of policy changes into price effects), having different values of $\eta_i(j)$ across states also have important policy implications. For instance, an expansionary (or a contractionary) monetary policy should affect the prices in each state individually rather than commonly (i.e., the case in which $\eta_i(j) = \eta(j)$ for all i) across states. Similarly, a fiscal policy (either at the country or state level) determining the tax rates would again affect the prices in each state individually rather than commonly. The different values of $\eta_i(j)$ across industries are also important for industry specific policies; e.g., a bailout plan to rescue a specific industry from a financial crisis should be formed completely different than rescuing another one in terms of determining its price setting behavior.

Robustness of the First-Step Analysis

The empirical results of the first-step analysis need further investigation for the possibility of having region specific consumption shares of industry j that are produced in the home country (i.e., $\gamma_i(j)$'s). So far, due to lack of accurate international trade data at the state level, it has been imposed that $\gamma_i(j) = \gamma(j)$ for each industry. However, if there are deviations from the national average of $\gamma(j)$ for any state, the calculated $\eta_i(j)$'s may be biased. In order to show this, by the help of Equation 2.4, consider the modified versions of Equations 4.1 and 4.3:

$$\alpha^H(j) \sum_r P^H_{r,r}(j) Y^H_r(j) = \sum_r \sum_i \left(\begin{array}{c} \dfrac{P^H_{r,r}(j)^{1-\left(\eta_i(j)+\left(\widetilde{\eta_i(j)}-\eta_i(j)\right)\right)} \left(\left(D^H_{i,r}\right)^{\delta(j)}\right)^{-\left(\eta_i(j)+\left(\widetilde{\eta_i(j)}-\eta_i(j)\right)\right)}}{\sum_m \left(P^H_{m,m}(j)\left(D^H_{i,m}\right)^{\delta(j)}\right)^{1-\left(\eta_i(j)+\left(\widetilde{\eta_i(j)}-\eta_i(j)\right)\right)}} \\ \\ \times \left(\gamma(j) + \left(\gamma_i(j) - \gamma(j)\right)\right) P_i(j) C_i(j) \end{array} \right) \tag{4.4}$$

and

$$\pi^H_r(j) = \frac{\sum_i \left(\eta_i(j) + \left(\widetilde{\eta_i(j)} - \eta_i(j)\right)\right) \left(\dfrac{\left(P^H_{r,r}(j)\left(D^H_{i,r}\right)^{\delta(j)}\right)^{-\left(\eta_i(j)+\left(\widetilde{\eta_i(j)}-\eta_i(j)\right)\right)} \left(\gamma(j)+\left(\gamma_i(j)-\gamma(j)\right)\right) P_i(j) C_i(j)}{\sum_m \left(P^H_{m,m}(j)\left(D^H_{i,m}\right)^{\delta(j)}\right)^{1-\left(\eta_i(j)+\left(\widetilde{\eta_i(j)}-\eta_i(j)\right)\right)}} \right)}{\sum_i \left(\left(\eta_i(j) + \left(\widetilde{\eta_i(j)} - \eta_i(j)\right)\right) - 1\right) \left(\dfrac{\left(P^H_{r,r}(j)\left(D^H_{i,r}\right)^{\delta(j)}\right)^{-\left(\eta_i(j)+\left(\widetilde{\eta_i(j)}-\eta_i(j)\right)\right)} \left(\gamma(j)+\left(\gamma_i(j)-\gamma(j)\right)\right) P_i(j) C_i(j)}{\sum_m \left(P^H_{m,m}(j)\left(D^H_{i,m}\right)^{\delta(j)}\right)^{1-\left(\eta_i(j)+\left(\widetilde{\eta_i(j)}-\eta_i(j)\right)\right)}} \right)} \tag{4.5}$$

where $\widetilde{\eta_i(j)}$ is the true value of the elasticity of substitution, and $\left(\widetilde{\eta_i(j)} - \eta_i(j)\right)$ is the bias in the elasticity of substitution due to having state-specific $\gamma_i(j)$'s. As is evident by these equations, non-zero $(\gamma_i(j) - \gamma(j))$'s may in turn lead to having non-zero $\left(\widetilde{\eta_i(j)} - \eta_i(j)\right)$'s. In a special case in which $\gamma_i(j) = \gamma(j)$ for all i, $\widetilde{\eta_i(j)} = \eta_i(j)$, which is the case in the empirical analysis, above. However, when $\gamma_i(j) \neq \gamma(j)$ for any i, $\widetilde{\eta_i(j)} \neq \eta_i(j)$, and thus, $\eta_i(j)$'s may be biased. In order to investigate the relation between $\left(\widetilde{\eta_i(j)} - \eta_i(j)\right)$'s and $(\gamma_i(j) - \gamma(j))$'s, the following question is asked: if the true values of the elasticity of substitution, $\widetilde{\eta_i(j)}$'s, are, in fact, 1% higher than their calculated values, $\eta_i(j)$'s $\left(\text{i.e., if } \frac{\widetilde{\eta_i(j)}-\eta_i(j)}{\eta_i(j)} = 0.01 \text{ for any } i\right)$, what would be the corresponding deviation of $\gamma_i(j)$'s from their national average $\left(\text{i.e., what would be } \frac{\gamma_i(j)-\gamma(j)}{\gamma(j)} \text{ for any } i\right)$? By using the available data, the median deviation of $\gamma_i(j)$'s, where median is taken across states, is calculated as -0.29% for food, 0.32% for apparel, 1.41% for electronics, and -0.23% for furniture, all leading to 1% of bias in calculated $\eta_i(j)$'s.[16] In other words, if

[16] The 25th (respectively, 75th) percentile deviation of $\gamma_i(j)$'s, where percentile is taken across states, is calculated as -0.13% (respectively, -0.36%) for food, 0.32% (respectively, 0.32%) for apparel, 1.40% (respectively, 1.41%) for electronics, and -0.21% (respectively, -0.27%) for furniture, all leading to 1% of bias in calculated $\eta_i(j)$'s.

$\gamma_i(j) > \gamma(j)$ for any i (i.e., if a state is consuming more domestic products compared to the national average), the calculated $\eta_i(j)$'s can be undervalued for food and furniture and overvalued for apparel and (especially) electronics, on average. According to the regression results in Table 3, if $\gamma_i(j) \neq \gamma(j)$ for any i, this result would not only support further the findings of this paper in terms of explaining the elasticities of substitution in a structural way, but also mean that consumption agglomerations are positively related to the consumption shares in all four industries that are produced in foreign countries; i.e., states with higher consumption agglomerations consume more international products (imports). In other words, the possibility of having region specific $\gamma_i(j)$'s not only supports the empirical findings of this paper related to the elasticities of substitution, but also provides further insight related to relation between international imports and consumption agglomerations.

4.2. Second-Step Analysis

Using the results of the first-step analysis for each industry (i.e., numerically solved 48 $\eta_i(j)$'s and one $\delta(j)$ for each state), in the second-step analysis, $\alpha_r^H(j)$ (for all r, j) are numerically solved using Equation 2.14 (where, for each industry, there are 48 unknown $\alpha_r^H(j)$'s and 48 versions of Equation 2.14, one for each state). As in the first-step analysis, using NLS, these 48 unknown $\alpha_r^H(j)$'s for each industry j are exactly identified via available data. Since empirically tested expression of this paper is again nonlinear (i.e., Equation 2.14), for each industry, the selection of the starting values in determining the NLS parameters (i.e., 48 $\alpha_r^H(j)$'s, one for each state r) are important. So, to be consistent with the available data of national level $\alpha^H(j)$ values, the starting value of $\alpha_r^H(j)$'s for food are set to 0.5969, for apparel set to 0.6267, and for electronics set to 0.0913, for furniture set to 0.6412.

Empirical Results of the Second-Step Analysis

Numerically calculated $\alpha_r^H(j)$ values are depicted in Figures 5-8 for food, apparel, electronics, and furniture, respectively. As is evident in Figure 5, most of the Western States (especially Mountain West) share higher $\alpha_r^H(j)$ values for food, while Midwestern and West South Central States (especially Texas and Arkansas), together with high GSP states such as California and Pennsylvania, share lower $\alpha_r^H(j)$ values. In other words, while most of the food produced low $\alpha_r^H(j)$ states are used as a final consumption good within the country, the food produced in high $\alpha_r^H(j)$ states are either used as an intermediate input or exported abroad.

According to Figure 6, except for high GSP states such as California, New York, Pennsylvania, and Texas, almost all states have higher $\alpha_r^H(j)$ values for apparel implying that the apparel production of most of these states are used as final good within the country. Only high GSP states such as California, New York, Pennsylvania, and Texas can produce intermediate inputs and export abroad.

Compared to food and apparel, there is a different story for electronics according to Figure 7: the $\alpha_r^H(j)$ values are low for most of the states implying that most of the electronics production is used either as intermediate input or exported abroad. Exceptions are some Western and Southern States.

Figure 8 depicts $\alpha_r^H(j)$ values for furniture. As is evident, except for East North Central, Middle Atlantic, and East South, together with high GSP states such as California and Texas, most of the states produce furniture that is consumed as a final good within the country.

A common feature of Figures 5-8 seems to be the negative relation between $\alpha_r^H(j)$ values and GSP levels of the states (especially in California, New York, and Texas). The correlation coefficients between $\alpha_r^H(j)$ values and other state-level variables are given in Table 4. Similar to the first-step analysis above, while the state-level industry-specific production and GSP are to capture the agglomeration effects, the state-level industry-specific production and export clusterings are to capture the specialization effects. In terms of production patterns, agglomeration effects are generally referred as economies of agglomeration which is generally credited to Alfred Marshall (e.g., see Krugman, 1991) and describes the benefits that firms obtain when locating near each other. It is related to the idea of economies of scale and network effects, in that the more related firms that are clustered together, the lower the cost of production (firms have competing multiple suppliers, greater specialization and division of labor) and the greater the market that the firm can sell into. Even when multiple firms in the very same sector (competitors) cluster, there may be advantages because that cluster attracts more suppliers and customers than a single firm could alone. In this context, economies of agglomeration may lead to lower values of $\alpha_r^H(j)$ which correspond to higher intermediate input production together with high international export. The intermediate input part of this story is consistent with Krugman and Venables (1995) and Venables (1996) who show that intermediate usage creates cost and demand linkages between firms and a tendency for manufacturing agglomeration. The international trade part of the story is consistent with Melitz (2003) who shows how the exposure to trade induces only the more productive firms (e.g., firms that benefit from economies of agglomeration) to enter the international export market, while some less productive firms continue to produce only for the domestic market. High negative correlation coefficients between $\alpha_r^H(j)$ values and industry-specific production and GSP in Table 4 support this story of economies of agglomeration for production. As is also evident, there is a negative relation between $\alpha_r^H(j)$ values and industry-specific production clustering and industry-specific export clustering. Since specialization and agglomeration are already correlated to each other, this is already related to the agglomeration effects as mentioned above. Moreover, the results related to specialization are consistent with the analysis of Amiti (1999) who shows that intermediate-good intensity has a positive and significant effect on geographical concentration.

As in the first-step analysis, analyzing the correlation coefficients does not depict the exact relation between $\alpha_r^H(j)$ values and agglomeration and specialization. Also, it is hard to make a comparison across industries with only correlation coefficients. Moreover, agglomeration effects can be correlated to specialization effects. In this context, a formal regression analysis is employed including final good usage, industry-specific production, and industry-specific production clustering. As is evident in Table 5, both independent variables have negative and significant effects on final good share; i.e., the products of highly specialized and agglomerated industries are used more as intermediate inputs or exported abroad. Across states of the U.S., 1 percent increase in agglomeration or specialization corresponds to around 0.3 to 0.4 percent fall in $\alpha_r^H(j)$ values for food, apparel, and furniture, while it corresponds to around 1 to 1.4 percent fall in $\alpha_r^H(j)$ values for electronics; i.e., agglomeration and specialization effects in electronics are around three times higher than other industries. This result shows the importance of information spillover across firms in the production of electronics. When both agglomeration and specialization effects are included in the regression analysis, only agglomeration effect becomes significant; as in the first-step analysis, this may be due to a possible multicollinearity between these independent variables. Finally, high explanatory power of the regressions again support the analysis.

In sum, the portion of production that is used as final good within the country (i.e., $\alpha_r^H(j)$'s) differ substantially across states of the U.S., and these differences can be systematically explained by the structure of the products together with the distribution of industry-specific production within the country.

5. Conclusions

This paper has introduced a model that relates consumption, production, and trade patterns of a region to location of all regions, income level of all regions, price level of all regions, as well as the good specific transportation costs, region/good specific technology levels, and factor costs. A couple of nuances are important to note in the model: (i) by assigning different elasticities of substitution across regions/firms, region/firm specific mark-up differences are allowed, (ii) the problematic iceberg assumption is avoided by employing more realistic trade-distance good-specific elasticities through a transportation sector, (iii) the portion of production that is used as a final good within the country is captured by firm/region specific parameters, (iv) international trade is controlled for by firm/region specific parameters.

The model has been numerically solved by state-level consumption and production data belonging to industries of food, apparel, electronics, and furniture from the U.S. The obtained parameters are further compared and connected to agglomeration and specialization of the industries in terms of both consumption and production. In particular, on the consumption side, it has been shown that the industry- and state-level elasticities of substitution can be significantly explained by consumption agglomerations; the elasticities are positively (respectively, negatively) affected by agglomeration of consumption for food and furniture (respectively, for apparel and electronics).

The differences across these industries are connected to the homogeneity of the products, where homogeneity is further supported by numerically calculated median elasticities of substitution across states/industries. Consumption agglomerations may also be connected to international imports at the industry and state levels. On the production side, it has been shown that the industry- and state-level portion of production that is used as a final good within the country can be significantly explained by both agglomeration and specialization of the industries; these portions are negatively related to both effects. In other words, the industry- and state-level portion of production that is used as an intermediate input or exported abroad is significantly and positively related to agglomeration and specialization of the industries across states. Thus, agglomeration and specialization of industries play an important role in determining the patterns of trade, both intranationally and internationally. Finally, comparisons across industries suggest that the spillover effects are much higher for electronics compared to food, apparel, or furniture, in terms of both consumption and production. High explanatory powers in the regression analyses further support the model.

An obvious next step is to investigate patterns of production, consumption, and trade by moving the analysis of this paper to an international context. In such a case, cross-country income differences or trade policies related to setting the optimal tariff rates can be shed more light through agglomeration and specialization effects. Alternatively, having the empirical results of this paper, parameters of the model related to intermediate input or international trade may be endogeneized through location theories. Such an analysis would have important policy implications in terms of determining the causality between trade and the distribution of economic activity, both intranationally and internationally.

References

[1] Alvarez, Fernando and Lucas, Robert Jr., (2007), "General equilibrium analysis of the Eaton-Kortum model of international trade," *Journal of Monetary Economics*, 54(6): 1726-1768.

[2] Anderson, J.E., (1979), "A Theoretical Foundation for the Gravity Equation", *American Economic Review*, 69(1): 106-116.

[3] Anderson, J.E., and Wincoop, E.V., (2003), "Gravity with Gravitas: A Solution to the Border Puzzle", *American Economic Review*, 93(1): 170-192.

[4] Anderson, J.E., and Wincoop, E.V., (2004), "Trade Costs", *Journal of Economic Literature*, 42: 691-751.

[5] Bergstrand, J.H., (1985), "The Gravity Equation in International Trade: Some Microeconomic Foundations and Empirical Evidence", *The Review of Economics and Statistics*, 67(3): 474-481.

[6] Bergstrand, J.H., (1989), "The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory in International Trade", *The Review of Economics and Statistics*, 71(1): 143-153.

[7] Berry, W.D., Fording, R.C., and Hanson, R.L., (2003), "Cost of Living Index for the American States, 1960-2003", Inter-University Consortium for Political and Social Research (ICPSR) Study No: 1275.

[8] Bowen, H.P., Leamer, E.E., and Sveikauskas, L., (1987), "Multicountry, Multifactor Tests of the Factor Abundance Theory", *American Economic Review*, 77(5): 791-809.

[9] Corsetti, G., and Pesenti, P., (2005), "International Dimensions of Optimal Monetary Policy", *Journal of Monetary Economics*, 52: 281-305.

[10] Deardorff, A. V. (1984): “Testing Trade Theories and Predicting Trade Flows,” in Handbook of International Economics, Volume I, ed. by R. Jones and P. Kenen. Amsterdam: North-Holland.

[11] Dornbusch, R., Fischer, S., and Samuelson, P.A., (1977), "Comparative Advantage, Trade, and Payments in a Ricardian Model with a Continuum of Goods", *The American Economic Review*, 67(5): 823-839. .

[12] Dornbusch, R., Fischer, S., and Samuelson, P.A., (1980), "Heckscher-Ohlin Trade Theory with a Continuum of Goods", *Quarterly Journal of Economics*, 95(2): 203-224.

[13] Eaton, J., and Kortum, S., (2002), "Technology, Geography and Trade", *Econometrica*, 70(5): 1741-1779.

[14] Erceg, C.J., Henderson, D.W., and Levin A.T., (2000), "Optimal Monetary Policy with Staggered Wage and Price Contracts", *Journal of Monetary Economics*, 46: 281-313.

[15] Evenett, S. J., and W. Keller (2002): “On Theories Explaining the Success of the Gravity Equation,” *Journal of Political Economy*, 110, 281–316.

[16] Fujita, M. and Thisse, J.F., 1996, "Economics of Agglomeration," *Journal of the Japanese and International Economies*, 10(4): 339-378.

[17] Gali, J. and Monacelli, T., (2005), "Monetary Policy and Exchange Rate Volatility in a Small Open Economy", *Review of Economic Studies*, 72: 707-734.

[18] Gallaway, Michael P. , McDaniel, Christine A., and Rivera, Sandra A., (2003), "Short-run and long-run industry-level estimates of U.S. Armington elasticities", *North American Journal of Economics and Finance* 14: 49–68

[19] Grossman, G. M., and E. Helpman (1995): “Technology and Trade,” in The Handbook of International Economics, Volume III, ed. by G. M. Grossman and K. Rogoff. Amsterdam: North-Holland.

[20] Helpman, E. (1987): “Imperfect Competition and International Trade: Evidence from Fourteen Industrial Countries,” *Journal of Japanese and International Economics*, 1: 62–81.

[21] Hillberry, R., and Hummels D., (2003), "Intranational Home Bias: Some Explanations", *The Review of Economic and Statistics*, 85(4):1089-1092.

[22] Hummels, D., (2001), "Toward a Geography of Trade Costs", *mimeo.*

[23] Hummels, D., and J. Levinsohn (1995):“Monopolistic Competition and International Trade: Reconsidering the Evidence,” *Quarterly Journal of Economics*, 110: 799–836.

[24] Hummels, D., D. Rapoport, and K. Yi (1998): “Vertical Specialization and the Changing Nature of World Trade,” *FRBNY Economic Policy Review*, 4: 79–99.

[25] Jensen, P. E. (2000): “Analysis of Bilateral Trade Patterns with Panel Data,” *Review of International Economics*, 8: 86–99.

[26] Krugman, P., (1980), "Scale Economies, Product Differentiation, and the Pattern of Trade", *American Economic Review,* 70(5): 950-959.

[27] Krugman, P., (1991), "Increasing Returns and Economic Geography", *Journal of Political Economy*, 99(3): 483-499.

[28] Krugman, P. R. (1995): “Increasing Returns, Imperfect Competition, and the Positive Theory of International Trade,” in The Handbook of International Economics, Volume III, ed. by G. M. Grossman and K. Rogoff. Amsterdam: North-Holland.

[29] Krugman, P. R., and A. J. Venables (1995): “Globalization and the Inequality of Nations,” *Quarterly Journal of Economics*, 110: 857–880.

[30] Lopez, Rigoberto A., Pagoulatos, Emilio, Gonzalez, Maria A., (2006), "Home bias and U.S. imports of processed food products", *North American Journal of Economics and Finance* 17: 363–373

[31] Matsuyama, K., (2000), "A Ricardian Model with a Continuum of Goods under Nonhomothetic Preferences: Demand Complementarities, Income Distribution, and North-South Trade", *Journal of Political Economy,* 108(6): 1093-1120.

[32] McCann, P., 1998, The Economics of Industrial Location: A Logistics-Costs Approach, 1998, Springer, Heidelberg.

[33] McCann, P., 2001, "A Proof of the Relationship between Optimal Vehicle Size, Haulage Length and the Structure of Distance-Transport Costs", Transportation Research A, 35.8, 671-693.

[34] McCann, P., (2005), "Transport costs and new economic geography", *Journal of Economic Geography*, 5: 305–318.

[35] McCann, P., and Fingleton, B., 1996, "The Regional Agglomeration Impact of Just-In-Time Input Linkages: Evidence from the Scottish Electronics Industry", Scottish Journal of Political Economy, 43.5, 493-518.

[36] Melitz, M.J., 2003, "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity", *Econometrica*, 71: 1695-1725.

[37] Parr, J.B., G.J.D. Hewings, J. Sohn, and S. Nazara, (2002), "Agglomeration and trade:
Some additional perspectives", Regional Studies, 36, 675-684.

[38] Rauch, J. E. (1999): "Networks vs. Markets in International Trade," *Journal of International Economics*, 48, 7–35.

[39] Rossi-Hansberg, E., (2005), "A Spatial Theory of Trade", *American Economic Review*, 95(5): 1464-1491.

[40] Samuelson, Paul A., (1954), "The Transfer Problem and Transport Costs, II: Analysis of Effects of Trade Impediments", *The Economic Journal*, 64(254): 264-289

[41] Suga, Nobuhito, (2007), "A monopolistic-competition model of international trade with external economies of scale", *North American Journal of Economics and Finance* 18: 77–91

[42] Trefler, D., (1995), "The Case of the Missing Trade and Other Mysteries", *American Economic Review*, 85(5): 1029-46.

[43] Wei, SJ, (1996), "Intranational versus International Trade: How Stubborn are Nations in Global Integration?", NBER Working Paper No: 5531.

[44] Wolf, H., (2000), ""Intra-National Home Bias in Trade", *Review of Economics and Statistics*, 82(4): 555-563.

[45] Yi, K. (2003): "Can Vertical Specialization Explain the Growth of World Trade?", *Journal of Political Economy*, 111(1): 52-102

[46] Yilmazkuday, H., (2008), "How Important is Technology: A Counterfactual Analysis". Available at SSRN: http://ssrn.com/abstract=995540.

[47] Yilmazkuday, H., (2009), "Understanding Interstate Trade Patterns". Available at SSRN: http://ssrn.com/abstract=995542.

Table 1 - Summary Results of First-Step Analysis

	Food	Apparel	Electronics	Furniture
$\eta_i(j)$	2.637	2.128	2.585	2.870
$\delta(j)$	0.021	0.001	0.057	0.033

Notes: Nonlinear Least Squares has been used as a numerical solution method. The median value of $\eta_i(j)$'s (where median is calculated across states) for each industry is given. The complete vector of $\eta_i(j)$'s that include state specific measures for each industry are available upon request.

Table 2 - Correlation of the Vector of $\eta_i(j)$'s with State-Level Variables

	Food	Apparel	Electronics	Furniture
Industry-Specific Consumption	0.71	−0.47	−0.52	0.63
Industry-Specific Consumption Clustering	−0.07	−0.34	−0.24	0.20

Notes: For each industry j (where j represents food, apparel, electronics, or furniture), Industry-Specific Consumption corresponds to the vector consisting of $P_i^H(j)\,C_i^H(j)$'s for all i, and Industry-Specific Consumption Clustering corresponds to the vector consisting of $\left(\frac{P_i^H(j)C_i^H(j)}{\sum_i P_i^H(j)C_i^H(j)}\right) \Big/ \left(\frac{\sum_j P_{i,i}^H(j)Y_i^H(j)}{\sum_i \sum_j P_{i,i}^H(j)Y_i^H(j)}\right)$'s for all i.

Table 3 - Regressions on $\eta_i(j)$'s

	Dependent Variable: Log $\eta_i(j)$		
	Equation (1)	Equation (2)	Equation (3)
Food			
Log Industry-Specific Consumption	0.005		0.006
	(6.90)		(1.38)
Log Industry-Specific Consumption Clustering		−0.002	−0.001
		(0.36)	(0.30)
R-bar sqd.	0.51	0.49	0.50
Apparel			
Log Industry-Specific Consumption	−0.009		−0.021
	(3.53)		(1.62)
Log Industry-Specific Consumption Clustering		−0.031	0.013
		(2.35)	(0.94)
R-bar sqd.	0.22	0.19	0.22
Electronics			
Log Industry-Specific Consumption	−0.065		−0.074
	(8.04)		(1.68)
Log Industry-Specific Consumption Clustering		−0.138	0.010
		(2.18)	(0.21)
R-bar sqd.	0.58	0.56	0.58
Furniture			
Log Industry-Specific Consumption	0.007		0.013
	(5.79)		(1.49)
Log Industry-Specific Consumption Clustering		0.016	−0.006
		(1.38)	(0.68)
R-bar sqd.	0.42	0.40	0.42

Notes: For each industry j (where j represents food, apparel, electronics, or furniture), Industry-Specific Consumption corresponds to the vector consisting of $P_i^H(j)\,C_i^H(j)$'s for all i, and Industry-Specific Consumption Clustering corresponds to the vector consisting of $\left(\frac{P_i^H(j)C_i^H(j)}{\sum_i P_i^H(j)C_i^H(j)}\right) \Big/ \left(\frac{\sum_j P_{i,i}^H(j)Y_i^H(j)}{\sum_i \sum_j P_{i,i}^H(j)Y_i^H(j)}\right)$'s for all i. T-statistics are in parenthesis. All data are demeaned for scale effects.

Table 4 - Correlation of the Vector of $\alpha_i^H(j)$'s with State-Level Variables

	Food	Apparel	Electronics	Furniture
Industry-Specific Production	-0.85	-0.82	-0.31	-0.90
Gross State Product	-0.61	-0.72	-0.45	-0.69
Industry-Specific Production Clustering	-0.46	-0.63	-0.17	-0.45
Industry-Specific Export Clustering	-0.49	-0.56	-0.14	-0.45

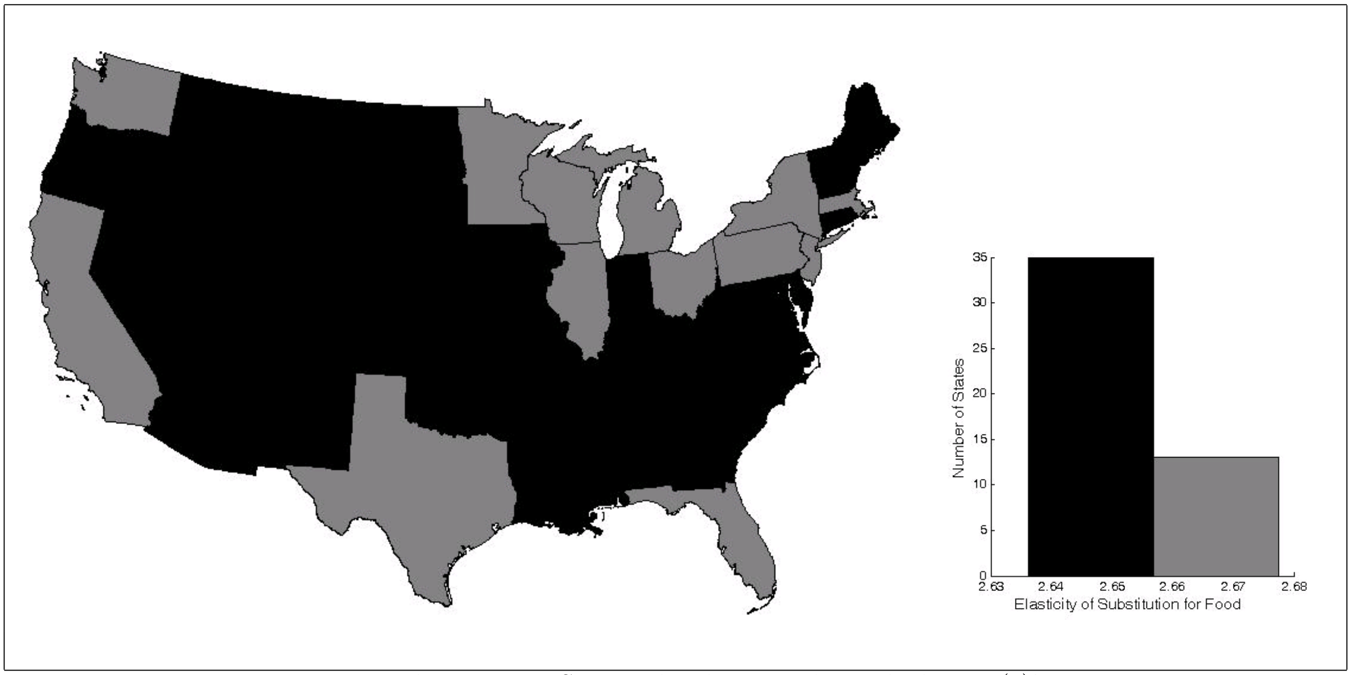
Notes: For each industry j (where j represents food, apparel, electronics, or furniture), Industry-Specific Production corresponds to the vector consisting of $P_{i,i}^H(j)\,Y_i^H(j)$'s for all i, Gross State Product corresponds to the vector consisting of $\sum_j P_{i,i}^H(j)\,Y_i^H(j)$'s for all i, Industry-Specific Production Clustering corresponds to the vector consisting of $\left(\frac{P_{i,i}^H(j)Y_i^H(j)}{\sum_i P_{i,i}^H(j)Y_i^H(j)}\right)\Big/\left(\frac{\sum_j P_{i,i}^H(j)Y_i^H(j)}{\sum_i\sum_j P_{i,i}^H(j)Y_i^H(j)}\right)$'s for all i, Industry-Specific Consumption Clustering corresponds to the vector consisting of $\left(\frac{P_i^H(j)C_i^H(j)}{\sum_i P_i^H(j)C_i^H(j)}\right)\Big/\left(\frac{\sum_j P_{i,i}^H(j)Y_i^H(j)}{\sum_i\sum_j P_{i,i}^H(j)Y_i^H(j)}\right)$'s for all i, Industry-Specific Export Clustering corresponds to the vector consisting of $\left(\frac{P_{i,i}^H(j)Y_i^H(j)}{\sum_i P_{i,i}^H(j)Y_i^H(j)}\right)\Big/\left(\frac{P_i^H(j)C_i^H(j)}{\sum_i P_i^H(j)C_i^H(j)}\right)$'s.

Table 5 - Regressions on $\alpha_i^H(j)$'s

	Dependent Variable: Log $\alpha_i^H(j)$		
	Equation (1)	Equation (2)	Equation (3)
Food			
Log Industry-Specific Consumption	-0.354		-0.351
	(9.33)		(7.64)
Log Industry-Specific Consumption Clustering		-0.342	-0.009
		(5.88)	(0.12)
R-bar sqd.	0.65	0.42	0.64
Apparel			
Log Industry-Specific Consumption	-0.304		-0.321
	(7.27)		(4.94)
Log Industry-Specific Consumption Clustering		-0.373	0.035
		(4.97)	(0.34)
R-bar sqd.	0.54	0.35	0.53
Electronics			
Log Industry-Specific Consumption	-0.978		-1.359
	(9.73)		(11.09)
Log Industry-Specific Consumption Clustering		-1.390	0.787
		(9.55)	(4.34)
R-bar sqd.	0.67	0.66	0.76
Furniture			
Log Industry-Specific Consumption	-0.301		-0.313
	(8.00)		(6.35)
Log Industry-Specific Consumption Clustering		-0.331	0.032
		(5.77)	(0.38)
R-bar sqd.	0.58	0.42	0.57

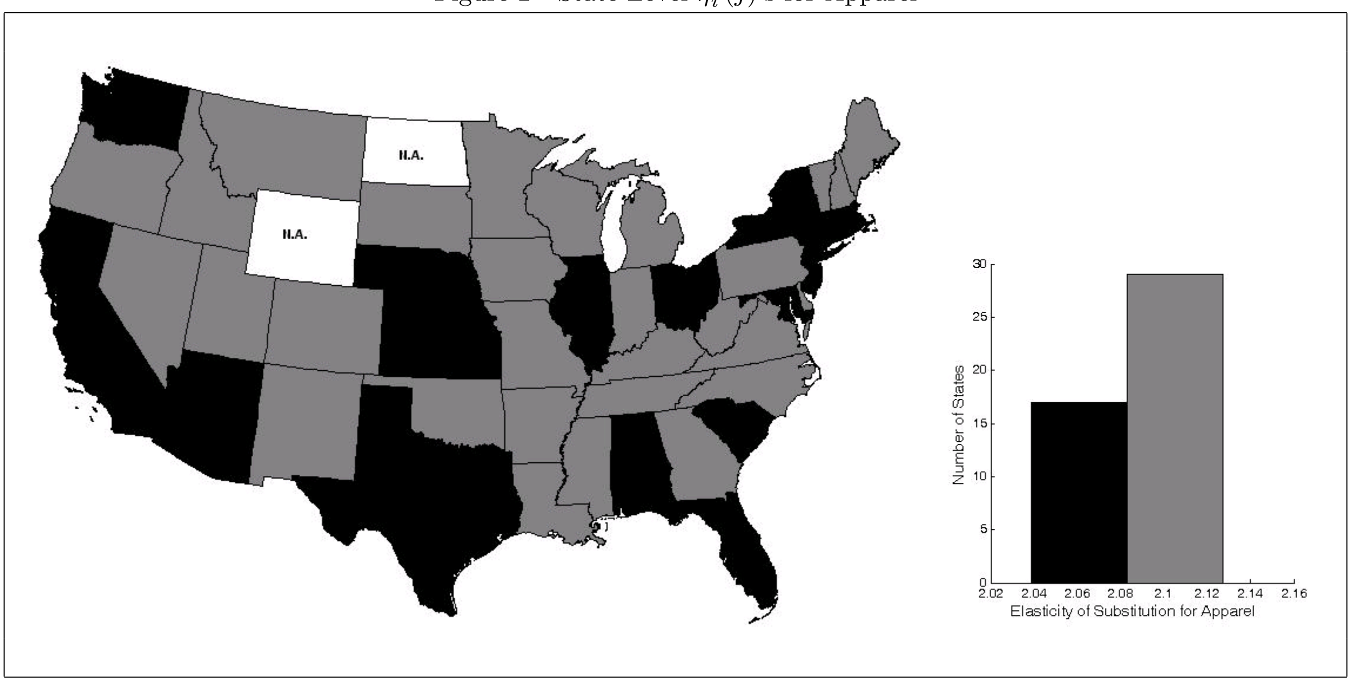
Notes: For each industry j (where j represents food, apparel, electronics, or furniture), Industry-Specific Consumption corresponds to the vector consisting of $P_i^H(j)\,C_i^H(j)$'s for all i, and Industry-Specific Consumption Clustering corresponds to the vector consisting of $\left(\frac{P_i^H(j)C_i^H(j)}{\sum_i P_i^H(j)C_i^H(j)}\right)\Big/\left(\frac{\sum_j P_{i,i}^H(j)Y_i^H(j)}{\sum_i\sum_j P_{i,i}^H(j)Y_i^H(j)}\right)$'s for all i. T-statistics are in parenthesis. All data are demeaned for scale effects.

Figure 1 - State-Level $\eta_i\,(j)$'s for Food



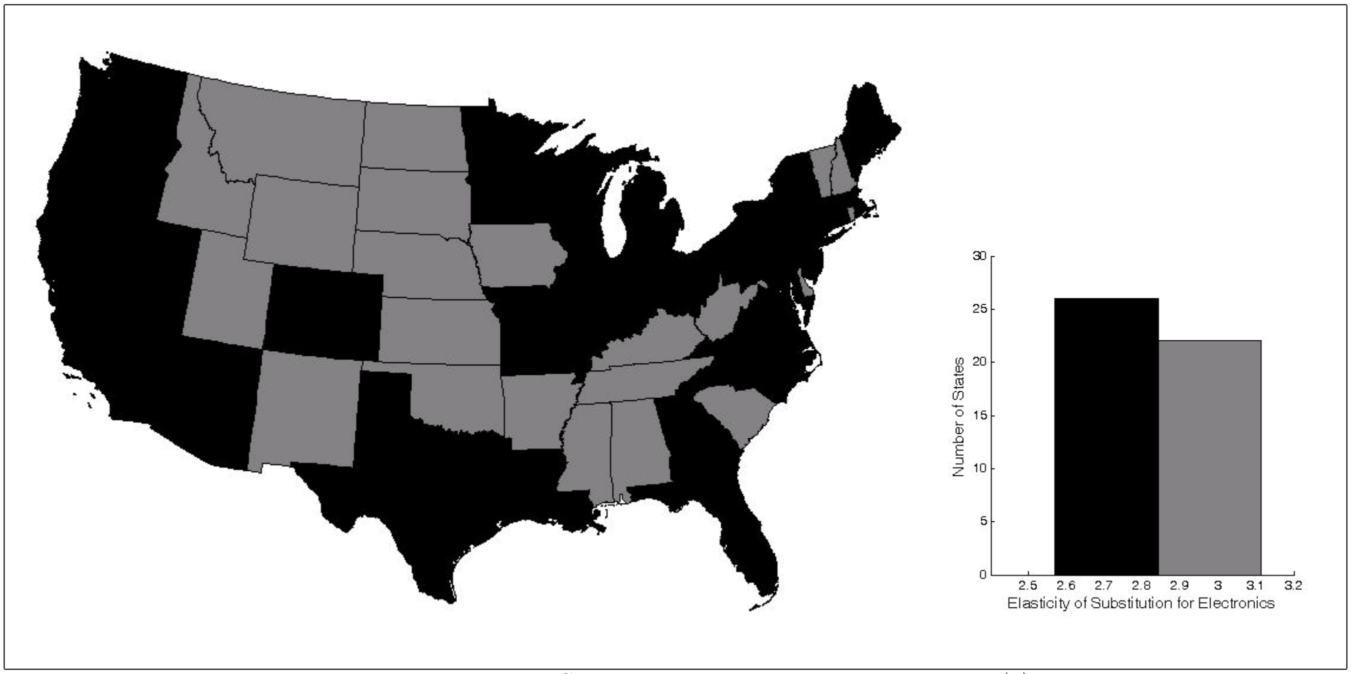
Notes: Nonlinear Least Squares has been used to calculate $\eta_i\,(j)$'s.

Figure 2 - State-Level $\eta_i\,(j)$'s for Apparel



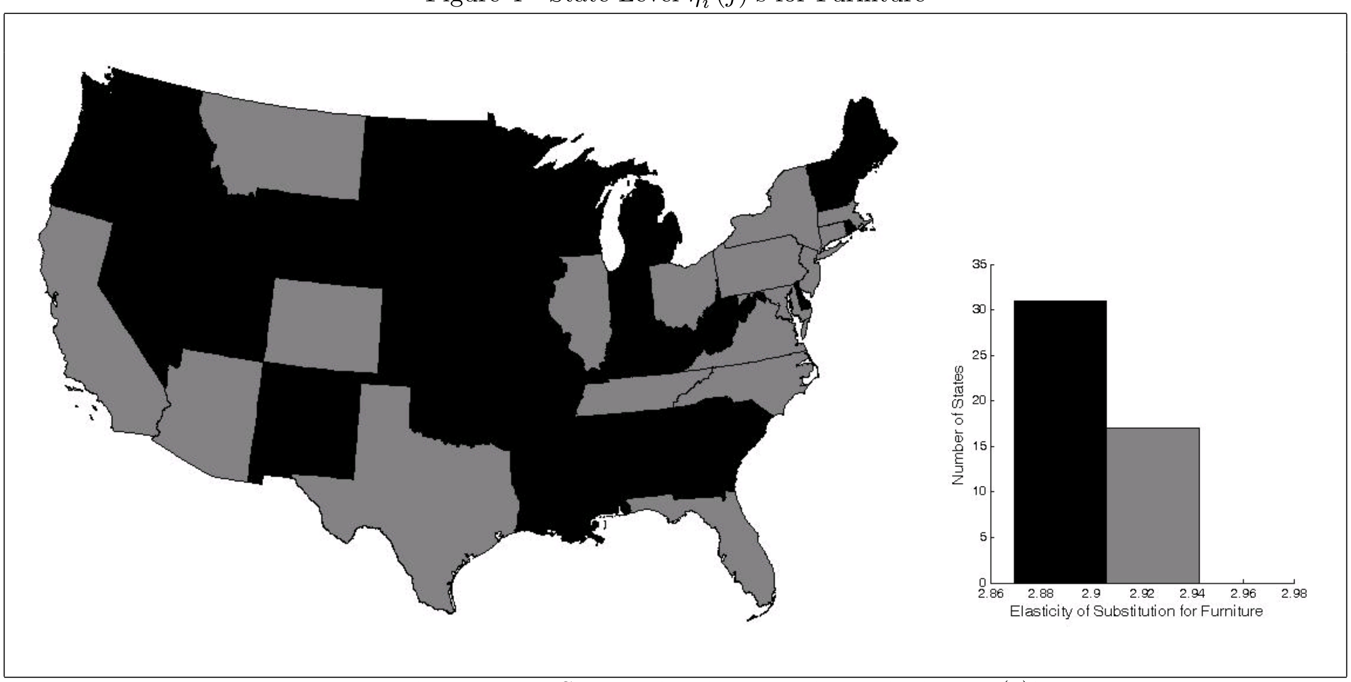
Notes: Nonlinear Least Squares has been used to calculate $\eta_i\,(j)$'s. There is no apparel production in North Dakota and Wyoming which are labelled as N.A..

Figure 3 - State-Level $\eta_i\,(j)$'s for Electronics



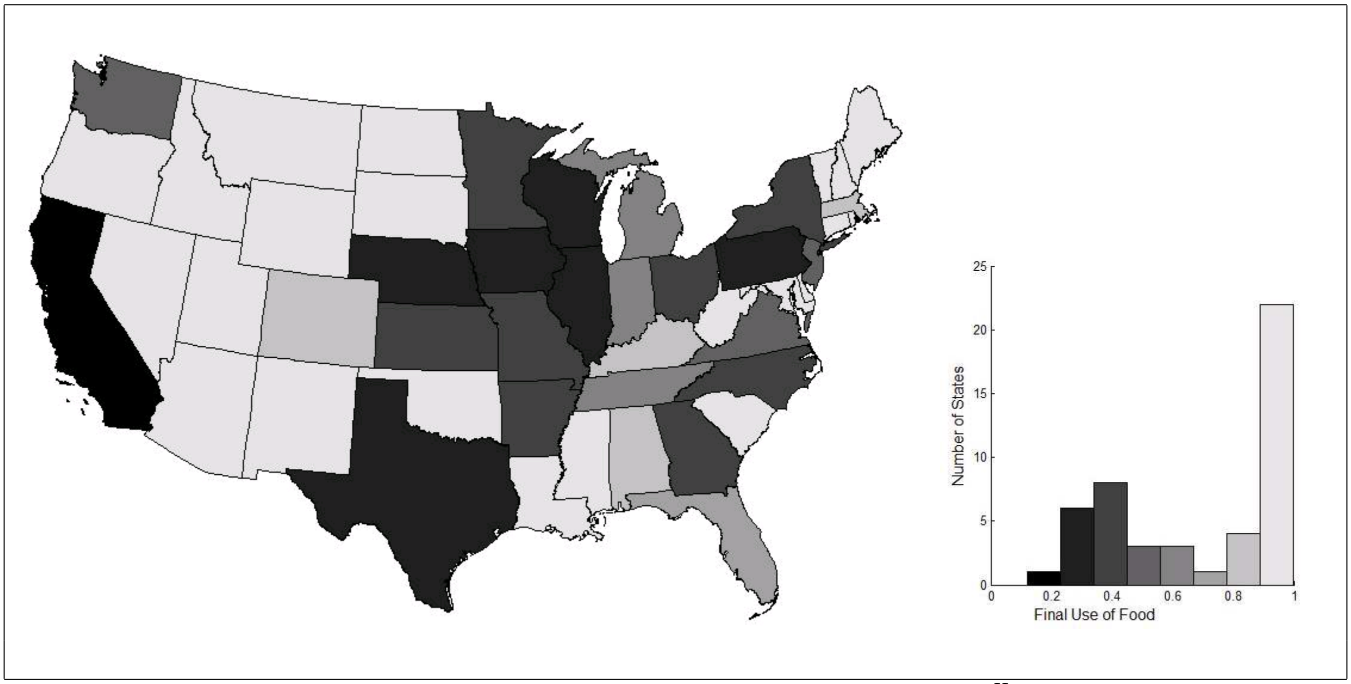
Notes: Nonlinear Least Squares has been used to calculate $\eta_i\,(j)$'s.

Figure 4 - State-Level $\eta_i\,(j)$'s for Furniture



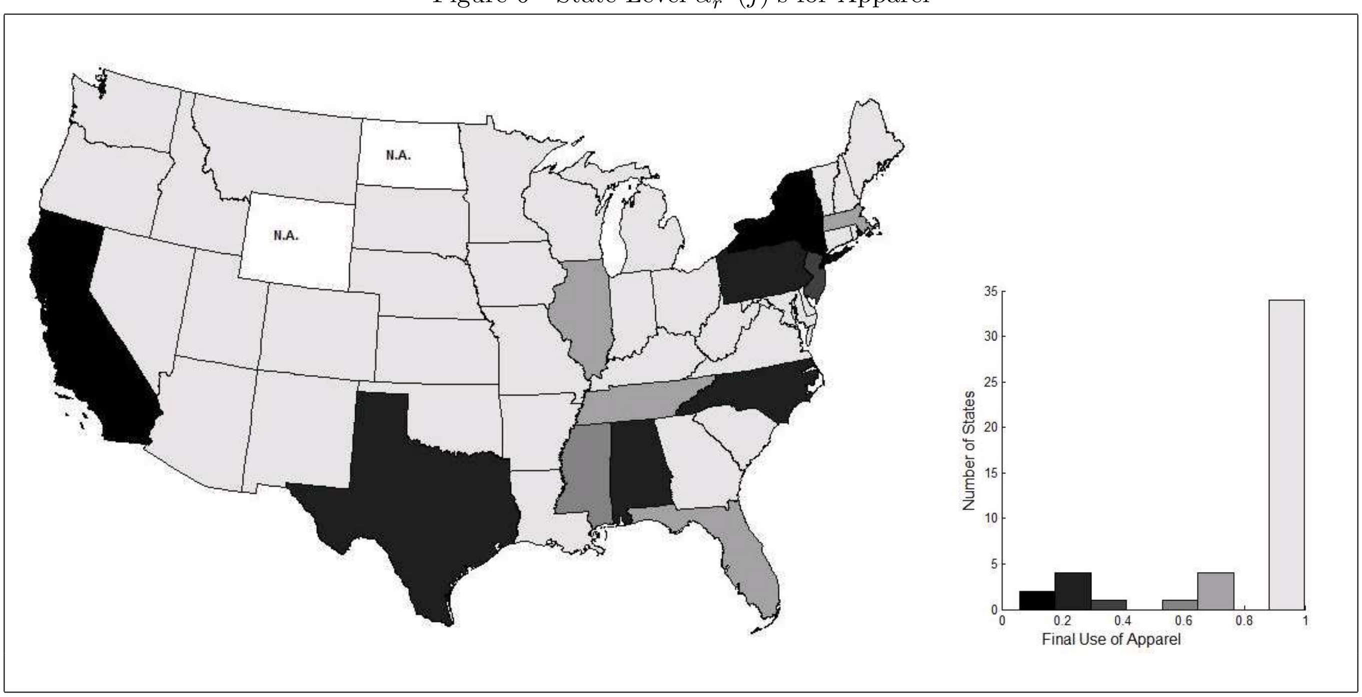
Notes: Nonlinear Least Squares has been used to calculate $\eta_i\,(j)$'s.

Figure 5 - State-Level $\alpha_r^H(j)$'s for Food



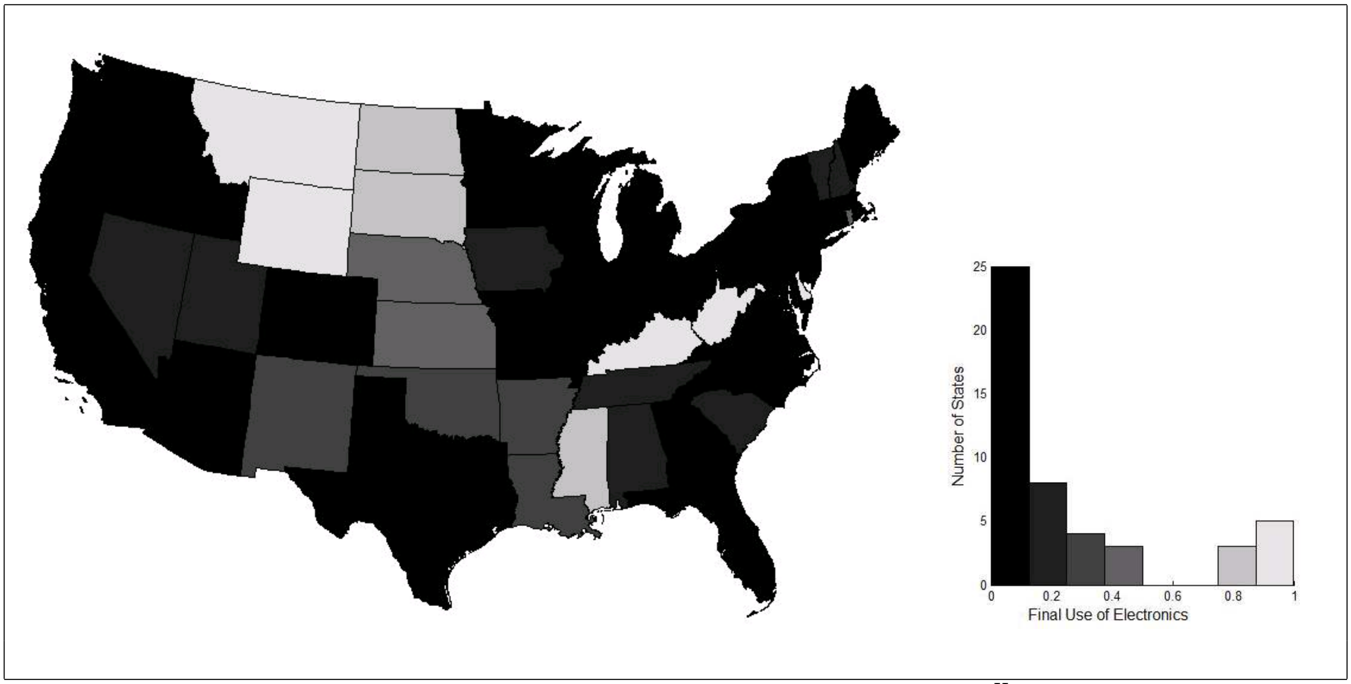
Notes: Nonlinear Least Squares has been used to calculate $\alpha_r^H(j)$.

Figure 6 - State-Level $\alpha_r^H(j)$'s for Apparel



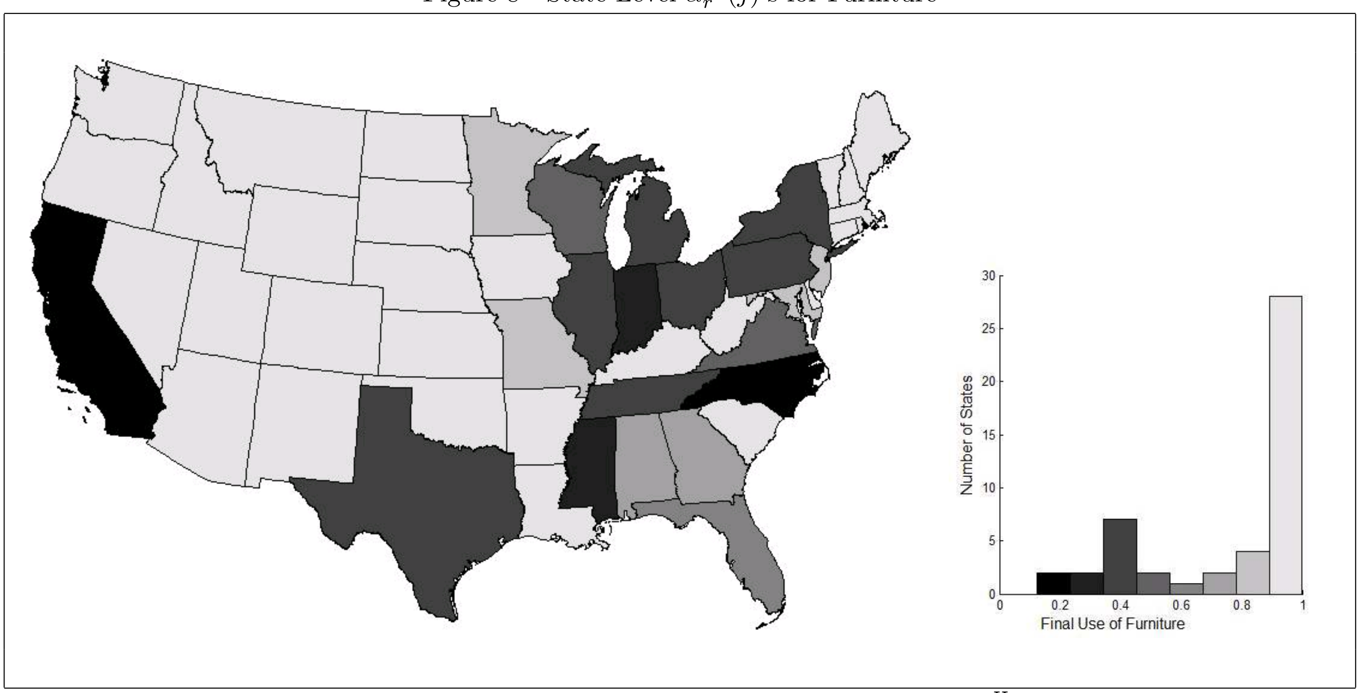
Notes: Nonlinear Least Squares has been used to calculate $\alpha_r^H(j)$. There is no apparel production in North Dakota and Wyoming which are labelled as N.A..

Figure 7 - State-Level $\alpha_r^H(j)$'s for Electronics



Notes: Nonlinear Least Squares has been used to calculate $\alpha_r^H(j)$.

Figure 8 - State-Level $\alpha_r^H(j)$'s for Furniture



Notes: Nonlinear Least Squares has been used to calculate $\alpha_r^H(j)$.