Specializing in Density:  
Spatial Sorting and the Pattern of Trade*  

Antoine Levy† and Jacob Moscona‡  
April 15, 2020

Abstract

This paper documents one way that domestic economic geography affects patterns of trade by showing that a country’s population distribution is an important source of comparative advantage. We develop a new strategy to estimate both the “population density affinity” of each industry and the “population concentration” of each country. We show that both US states and countries with more concentrated populations disproportionately export in sectors with high population density affinity. The findings are similar using an instrumental variables strategy in which we exploit variation in countries’ historical city size distribution to construct instruments for modern population concentration. We rationalize these findings with a model in which sector-specific exports are determined by the distribution of productivity within countries, and show how city-level data can be aggregated to measure determinants of country-level specialization. In the model, countries with higher population-weighted population density specialize in sectors that benefit most from agglomeration. Even conditional on aggregate endowments, our results suggest that the distribution of population within countries and the extent to which population is concentrated in dense cities shape comparative advantage.

JEL codes: F14, F16, R12, R13.

Keywords: International trade, comparative advantage, density, spatial sorting.

---

*We thank Daron Acemoglu, Treb Allen, David Atkin, Abhijit Banerjee, Arnaud Costinot, Dave Donaldson, Mayara Felix, Teresa Fort, Joel P. Flynn, Nathan Nunn, Ben Olken, Pierre-Luc Vautrey, Ivan Werning, and participants at the Urban Economics Association annual meetings, the MIT Development Lunch, and MIT Trade Tea for feedback and comments. First draft: November, 2018.

†Department of Economics, M.I.T., 50 Memorial Drive, Cambridge, M.A. 02142, (email: levya@mit.edu; website: http://economics.mit.edu/grad/levya)

‡Department of Economics, M.I.T., 50 Memorial Drive, Cambridge, M.A. 02142, (email: moscona@mit.edu; website: http://economics.mit.edu/grad/moscona)
1 Introduction

Does the distribution of economic activity within a country affect what it exports? Most analyses of patterns of trade treat countries as unified factor markets or equilibrium “points” in the production space. In this framework, countries are characterized by autarky supply and demand conditions that determine their ability to produce goods at competitive world prices. A growing body of research, however, documents an important interplay between within-country heterogeneity and cross-country trade.\(^1\) While existing work has highlighted the effects of trade on economic geography—the distribution of wages, employment, or economic activity—our focus is on the reverse relationship. This paper hypothesizes that variation across countries in domestic heterogeneity is an important determinant of patterns of cross-country trade.

While this idea is general and a version of it dates back to Courant and Deardorff (1992) and Courant and Deardorff (1993)’s “lumpiness” hypothesis, we take it to data by investigating one particular but central example. A major focus of recent work on economic geography is the observed variation across regions in population density, as well as how urban density differentially boosts the productivity of different industries.\(^2\) This logic suggests that the extent to which a country’s population is concentrated in cities might affect not only its domestic productivity, but also its international specialization and the composition of its exports.

If sectors benefit differently from population density, holding all other country-level characteristics constant, countries with a more concentrated population distribution will have a revealed comparative advantage in sectors that benefit disproportionately from agglomeration. Using newly developed measures of industry-level “density affinity” and country-level “population concentration,” we argue that urban density is a major determinant of not only the distribution of domestic economic activity, but also international exports.

We first present a model that illustrates how the distribution of factors of production within countries—i.e. having a concentrated versus dispersed population—affects patterns of trade. We assume that sectors vary in the extent to which they benefit from the population density of the location in which production takes place. In the baseline model, we are agnostic about the source of this variation in agglomeration externalities and simply assume that it exists. Countries are composed of cities that are endowed with different sector-neutral productivities. Endogenously, countries with more dispersion in sector-neutral productivity across cities exhibit higher population-weighted density. That is, they end up with a larger share of the population in dense cities.

The theoretical framework provides three key insights. First, motivated by evidence that countries display significant domestic spatial heterogeneity in factor prices, product specialization, and relative productivity (e.g. Porter, 2003; Desmet and Rossi-Hansberg, 2013), our model formalizes the idea that the relevant units of observation for understanding comparative advantage are regions

---

\(^1\)See, for example, Autor, Dorn, and Hanson (2013), Caliendo, Dvorkin, and Parro (2015), Dix-Carneiro and Kovak (2015), and Ramondo, Rodriguez-Clare, and Saborío-Rodríguez (2016), or the work in progress by Bakker (2018).

\(^2\)On the role of population density in economic geography and spatial sorting see Keesing and Sherk, 1971; Ciccone and Hall, 1996; Duranton and Puga, 2004; Moretti, 2012 and more recently Davis and Dingel, 2014 and Gaubert, 2018. On the impact of density on sector-specific productivity and role of density in determining heterogeneity across sectors in spatial sorting, see, for example, Nakamura, 1985; Rosenthal and Strange, 2004; Faggio, Silva, and Strange, 2017.
within countries where production takes place. This is different from most models of comparative advantage, which focus on aggregate country-level characteristics that are taken as given. Second, our model documents how regional data and characteristics can be aggregated to uncover country-level determinants of comparative advantage. For example, the model motivates our use of “population-weighted density” as the country-level summary of within-country heterogeneity in population density. Finally, the model provides theoretical justification for our main empirical framework and result: countries with higher population-weighted density have relatively lower autarky prices in sectors that benefit from agglomeration; hence, their exports exhibit a revealed comparative advantage in these sectors.

The rest of the paper empirically investigates whether the distribution of population within countries is an important determinant of comparative advantage. Our empirical strategy requires two main ingredients: (i) a sector-level causal estimate of “density affinity,” or the extent to which production in each sector is disproportionately located in denser locations, and (ii) a country-level estimate of population concentration.

To measure industry-level density affinity, we turn to detailed business location data across US urban areas from the County Business Patterns (CBP) and non-parametrically estimate the extent to which each sector is disproportionately located in denser locations. To account for potential endogeneity in the correlation between density and industry specialization, we use subterranean geological instruments that exogenously shift local density independently from other city-level characteristics. This procedure generates causal estimates of the marginal impact of a change in density on industry-level production. In the end, this procedure yields industry-level measures of density affinity across all 4-digit NAICS manufacturing sectors; the substantial heterogeneity in density affinity that we estimate lends credibility to the modeling assumption that there is significant variation in sector-specific sorting with respect to population concentration.

To measure population-weighted density across regions and countries, we rely on satellite-derived gridded population data from the LandScan database. The LandScan database incorporates comprehensive country-level census data on the distribution of population, and derives gridded population estimates using “smart interpolation,” a multi-layered, asymmetric, spatial modeling approach. These data make it possible to estimate characteristics of the geographic population distribution of each country. To measure population-weighted density, we sum population density across grid cells within each country, weighting each cell by its total population. This captures the experienced population density of the average person in the country.

Armed with these estimates, we investigate the relationship between density and comparative advantage. Before turning to cross-country trade, we focus on the exporting patterns of US States. Using the LandScan data, we estimate the population-weighted density of each state, and document that denser states indeed export relatively more in “density-loving” sectors. While this result is

---

3The data are available here: https://landscan.ornl.gov/landscan-datasets. We use the 2016 edition of the data set.
4For more information, see here https://landscan.ornl.gov/documentation
5While some recent studies have attempted to estimate export data at the metropolitan level (see e.g. the database constructed by (Tomer and Kane, 2014)), most trade flows data are still collected at a broader level of aggregation. The lowest level of consistent and exhaustive trade reporting in the United States is the state.
Figure 1: Population weighted population density across countries (deciles). The figure is a map in which countries are color-coded based on their population-weighted density decile. Darker countries have higher population-weighted density.

A preliminary test of our hypothesis, it also validates our density affinity estimates as supply side determinants of sector productivity rather than the product of path dependence, or demand-side forces. That is, our estimates of density affinity from the city-level US data could have been driven by the fact that, for historical or demand-side reasons, certain sectors are over-represented in certain US cities; this does not necessarily imply that there is a systematic relationship between density and sector-level productivity. However, the state-level export results suggests that density-loving sectors are indeed more productive in denser regions within the US.

Next, we investigate the role of density as a source of country-level comparative advantage. Country-level estimates of population weighted density are displayed in the map in Figure 1. Visually, there is substantial variation in density across countries, even within continents and income levels. For example, Finland and Sweden are two of the wealthiest and also two of the least dense countries in the world, by our measure; indeed, both countries have strong revealed comparative advantage in pulp and paper product exports, one of the least density-loving sectors. Within sub-Saharan Africa, Botswana is among the least dense countries while the nearby Democratic Republic of Congo and Djibouti, among the world’s poorest countries, are among the densest. Djibouti, meanwhile, exhibits a strong revealed comparative advantage in semiconductors, one of the most density-loving sectors. Finally, the United States has mid-range population-weighted density since it has both very dense cities, as well as a relatively large share of the population living in suburbs, towns, and rural areas. Anecdotally, these countries’ export patterns are consistent with our hypothesis.

---

We find that this pattern is systematic. Countries with higher population-weighted density have a revealed comparative advantage in density-loving sectors. This finding is robust to the inclusion of a broad range of country and industry-level controls, including the skill and capital intensity of each sector, as well as country-level income, skill endowment, specialization in agriculture, and several other controls that might bias the relationship between population-weighted density and a country’s composition of exports. The results also remain similar across a range of possible parameterizations of the key independent variable.\(^8\)

To correct for potential reverse causality from trade flows to density (see Krugman and Elizondo, 1996; Ades and Glaeser, 1995), we exploit differences in states’ and countries’ historical city sizes to construct instruments for modern variation in density. Data on the global distribution of cities and their populations for historical periods were collected by Chandler (1987), and recently digitized by Reba, Reitsma, and Seto (2016). While trade might affect modern economic geography, it is unlikely that modern patterns of trade, which have evolved substantially in recent decades and particularly after the Second World War, affected the historical (c. 1900) distribution of cities within countries.\(^9\) Using this identification strategy, the estimated effect of density on trade flows from our baseline results remains virtually unchanged. In our sample of countries, we find that the impact of the within-country population distribution on patterns of trade is comparable to and if anything slightly larger in magnitude than the impact of human or physical capital.

Finally, we investigate potential channels of causality underpinning the relationship density affinity and trade. While our “density affinity” estimates were non-parametric and reduced form, we document a similar pattern to our baseline results when we focus instead on the research and development (R&D) intensity of each industry, consistent with evidence that dense cities facilitate and spur innovation (Duranton and Puga, 2001; Duranton and Puga, 2004). We also document a similar pattern when we focus on the reliance of each industry on immobile natural resources, consistent with the idea that only industries that do not rely on natural resources are free to locate in cities (Ades and Glaeser, 1995). However, the combination of these channels do not fully explain our baseline results, suggesting that additional and unobservable industry-level agglomeration forces are also at play. We also rule out a set of additional possible channels, including skill and capital abundance, as well as the reliance of each industry on service-sector inputs.\(^10\)

Our key finding is that the distribution of population within countries is a key determinant of cross-country specialization and comparative advantage. This study is at the intersection of several broad areas of research. Our theory is most closely related to Courant and Deardorff (1992) and Courant and Deardorff (1993), who argue that patterns of trade come not only from relative factor abundance, but also from factor distribution (“lumpiness”). The potential relevance of “lumpiness” in relative factor endowments for regional specialization has been explored more recently by Debaere (2004), Bernard, Robertson, and Schott (2010), and Brakman and Van Marrewijk (2013). We

---

\(^8\)The results are also qualitatively identical using either Poisson pseudo maximum likelihood estimation or ordinary least squares.

\(^9\)For a detailed discussion of the evolution of US patterns of trade, see Irwin (2017).

\(^10\)On the relationship between service sector sourcing and economic geography, see Abdel-Rahman and Fujita (1990), Abdel-Rahman and Fujita (1993), Abdel-Rahman (1994), and Abdel-Rahman (1996)
contribute to this literature by using new satellite data to estimate variation across countries in population concentration—one form of factor “lumpiness”—and directly investigating its impact on comparative advantage.

This paper is also related to investigations of the sorting of sectors across cities (most recently Davis and Dingel, 2014; Gaubert, 2018). A range of work has documented that agglomeration benefits some sectors more than others and there is substantial heterogeneity in sector-specific sorting with respect to population density (Nakamura, 1985; Rosenthal and Strange, 2001; Rosenthal and Strange, 2004; Holmes and Stevens, 2004; Ellison, Glaeser, and Kerr, 2010; Faggio, Silva, and Strange, 2017). We extend existing work in this area by developing a reduced-form strategy to estimate industry-specific “density affinities” and, more importantly, investigating the relationship between domestic sorting of production and international trade.

Other work has focused on the impact of within-country trade costs on patterns of trade (Rauch, 1991; Coşar and Fajgelbaum, 2016). A large theoretical literature on patterns of trade arising from agglomeration, initiated by Krugman (1991), has given rise to studies of the stylized interaction between agglomeration and more traditional sources of comparative advantage (Van Marrewijk et al., 1997; Ricci, 1999; Pflüger and Tabuchi, 2016). We suggest that agglomeration may benefit sectors differentially even within manufacturing, and provide evidence that it shapes comparative advantage in cross-country trade.

Finally, our empirical framework builds on existing assessments of sources of comparative advantage across countries; recent empirical analyses that rely on a similar framework but investigate country-level characteristics include Nunn (2007), Chor (2010), Costinot (2009), Bombardini, Galipoli, and Pupato (2012), and Cingano and Pinotti (2016).

The paper is organized as follows. Section 2 provides a simple formalization of our hypothesis that comparative advantage across countries stems, in part, from the distribution of population within countries. Section 3 describes the data used in the empirical analysis. Section 4 presents our main results and Section 5 concludes.

2 Theoretical Framework

We present a model that illustrates how within-country heterogeneity in productivity can affect a country’s pattern of exports across industries. We emphasize how two key ingredients – (i) productivity heterogeneity across a country’s locations and (ii) differential returns to agglomeration across industries – can produce patterns of specialization both within and across countries. The theoretical results guide our estimation of the key components of our empirical analysis: industry-level “density affinity” and country-level “population-weighted density,” the notion of density that “makes sense” from the perspective of the model at the country level.

2.1 Environment: the closed economy

We study an economy where countries exhibit domestic heterogeneity across inhabited locations, or “cities.” A country is defined as a continuum of cities, indexed by $c \in C$, with innate productivity $A_c$. 
land area $B_c$, and equilibrium population $L_c$. The country’s total population is fixed to $L$; workers are mobile across regions within a country, but not across borders. The economy consists of $J$ tradable sectors indexed by $j = 1, \ldots, S$, as well as a non-tradable good specific to each city, “housing” ($H_c$); housing is the key force of pecuniary congestion in the model. Tradable goods can be shipped from city $c$ to city $d$, by paying iceberg trade costs $\tau_{c,d} \geq 1$, possibly equal to one (for within-country transactions).

2.1.1 Workers

Workers in city $c$ inelastically supply one unit of labor, earning wage $w_c$. They derive utility $U_c$ from the consumption of housing and a basket of tradable sectors:

$$U_c(h_c, c_j, j = 1, \ldots, J) = \left( \frac{h_c}{\beta} \left( \frac{\prod_{j=1}^S (\frac{c_j}{\bar{c}_j})^\alpha_j}{1 - \beta} \right) \right)^{1 - \beta}$$

where $h_c$ is the worker’s housing and $c_j$, total consumption of sector $j$, is a CES aggregate of a continuum of varieties indexed by $\omega$:

$$c_j = \left( \int_0^1 c_j(\omega)^{\frac{1}{1-\sigma}} d\omega \right)^{\frac{1}{1-\sigma}}$$

The price level in each sector $j$ is therefore: $p_j = \left( \int_0^1 p_j(\omega)^{1-\sigma} d\omega \right)^{\frac{1}{1-\sigma}}$, and the aggregate tradable price level in the country $P$ is: $P = \prod_{j=1}^J p_j^\alpha_j$. We assume that $\sigma > 1$, so that within each sector, varieties $\omega$ are substitutes. Indirect utility in city $c$ for a worker that supplies a unit of labor is thus: $V_c = \frac{Y_c}{p_{hc}^{1-\beta}}$. Since utility must be the same for a worker in all cities at some level $V_c = \bar{U} \forall c$, city-level income $Y_c$ is:

$$Y_c = \bar{U} P^{1-\beta} p_{hc}^\beta$$

As is standard, income is increasing in the price of housing $p_{hc}$.

2.1.2 Housing

The supply of land in location $c$ is fixed at $B_c$; this generates the key pecuniary congestion force in the model. As in Gaubert (2018), atomistic landowners in city $c$ own an amount $\gamma$ of local land, and produce housing using land and tradable goods, according to the production function:\footnote{For simplicity, we assume that they divide spending on final goods used as inputs in housing production across the $S$ sectors in the same manner as workers; alternatively, one could model the other input into housing production as migrant labor living at zero cost on rural land and only consuming the final good.}

$$H_c(\gamma) = \gamma^\xi \left( \frac{X_{hc}(\gamma)}{1 - \xi} \right)^{1-\xi}$$

Equalizing supply and demand yields equilibrium housing prices in each city:\footnote{The details are given in Appendix B.}

$$p_{hc}^{\frac{1}{\xi}} = \beta \frac{L_c Y_c}{B_c P_{hc}^{\frac{1}{\xi}}}$$

$$\bar{U}$$
All Ricardian rents accruing to local landowners are fully taxed by the city government and rebated to resident workers as lump-sum transfers $T_c$, as in Helpman (1998). Thus, disposable income $Y_c$ of a worker in city $c$ is proportional to wage income $w_c$: $Y_c = w_c + T_c = \frac{w_c}{1 - \beta \xi}$. Using the spatial equilibrium condition (2.1), we derive an expression for city-specific wages:

$$w_c = P(1 - \beta \xi) \hat{U}^{\frac{1}{\sigma - 1}} \beta^{\frac{\theta}{\sigma - 1}} \frac{\rho_c}{B_c} \frac{1}{1 - \beta \xi} \approx P \times D_c^{\frac{\theta}{\sigma - 1}}$$

where $D_c$ is the population density of city $c$. Consistent with a large literature in urban economics (Glaeser and Gottlieb, 2009), there is a log-linear relationship between local wages and local population density.

2.1.3 Production

To study the impact of density on industrial geography and trade, we turn to the supply side of the economy. For simplicity, labor $L_{jc}(\omega)$ is the only input to production. In each industry $j$, the output of variety $\omega$ in city $c$, $Q_{jc}(\omega)$, is given by:

$$Q_{jc}(\omega) = \bar{A}_{jc} L_{jc}(\omega)$$

Each city draws a Ricardian productivity parameter in each sector, $\bar{A}_{jc}$, from a Fréchet distribution. The unit cost of production for variety $\omega$ in sector $j$ and location $c$ is then $\frac{w_c}{\bar{A}}$. The actual productivity draw $\bar{A}_{cj}(\omega)$ for a variety of good $j$ in location $c$ has cumulative distribution function:

$$\text{Pr}(\bar{A}_{cj}(\omega) \leq \bar{A}) = F_{jc}(\bar{A}) = \exp\left(-\left(\frac{\bar{A}}{\bar{A}_{jc}}\right)^{\theta}\right)$$

Here we introduce the key assumption of the model, which allows us to isolate our channel of interest: the relationship between the population distribution and comparative advantage. We assume that a sector’s productivity in city $c$ depends on (i) the city’s exogenous sector neutral productivity term $A_c$, (ii) the city’s equilibrium population density $D_c$, and (iii) the extent to which each sector benefits from local density, $\eta_{jc}$. In particular, we let: $A_{jc} = A_c D_c^{\gamma_j}$. The sector-specific ”density elasticity” $\gamma_j$, mediates the relationship between density and sector-specific productivity. This implies that:

$$\text{Pr}(\bar{A}_{cj}(\omega) \leq \bar{A}) = \exp\left(-\left(\frac{\bar{A}}{A_c D_c^{\gamma_j}}\right)^{\theta}\right)$$

With this formulation, as long as all locations in a country are inhabited, the allocation of labor to any sector in any inhabited city of any country will never be exactly zero.

The variation in $\eta_j$ across sectors – the extent to which each sector benefits from local agglom-
eration – will be central to our empirical analysis, and is the key modeling assumption. The idea that industries benefit differentially from urban density is backed by substantial evidence (e.g. Nakamura, 1985; Rosenthal and Strange, 2004; Faggio, Silva, and Strange, 2017) and corroborated by our estimates in Section 3 below.\(^{15}\)

2.1.4 Trade across cities

If we make the (admittedly strong) assumption that trade costs are zero within country,—i.e. for two cities \(c_n\) and \(c_n'\) in country \(n\), \(\tau_{c_n c_n'} = 1\)—cost minimization by consumers in any location \(d\) then implies that the share of spending on varieties from location \(c\) in sector \(j\) must be equal for any locations \(d\) in the same country:\(^{16}\)

\[
\pi_{dcj} = \pi_{cj} = \frac{p_{cj}X_{dcj}}{X_{dj}} = \frac{(A_c D_{c}^j)^\theta w_c^{-\theta}}{\sum c'(A_c D_{c'}^j)^\theta w_{c'}^{-\theta}}
\]

(2.3)

where \(\pi_{dcj}\) denotes spending in city \(d\) on goods in sector \(j\) produced in city \(c\).

2.1.5 Equilibrium

Goods market clearing  In the equilibrium of the closed domestic economy, the wage bill in each sector \(j\) and city \(c\) equals total spending on goods produced in sector \(j\) in city \(c\).\(^{17}\) This generates the tradable goods market clearing condition:

\[
w_c L_{jc} = \alpha_j \frac{(A_c D_{c}^j)^\theta w_c^{-\theta}}{\sum c'(A_c D_{c'}^j)^\theta w_{c'}^{-\theta}} \sum d w_d L_d
\]

(2.4)

In the absence of within-country trade costs, the price index for good \(j\) is independent of the location where it is consumed and is proportional to:\(^{18}\)

\[
p_j \propto \left[ \sum c'(A_c D_{c'}^j)^\theta w_{c'}^{-\theta} \right]^{-\frac{1}{\theta}} \propto \left[ \sum c'(A_c D_{c'}^j)^{\theta - \frac{\beta \xi}{1 - \beta \xi}} \right]^{-\frac{1}{\theta}}
\]

(2.5)

Trade balance requires that tradable spending from all locations on all goods produced in location \(c\)

---

\(^{15}\)We remain agnostic here about the specific source of sector-specific density affinity; in section 4.5, we explore potential mechanisms, like a lesser reliance on natural resources, local sharing of non-tradable services inputs, or higher innovation intensity.

\(^{16}\)This is derived in Appendix B. This relies on standard Eaton-Kortum algebra similar to Costinot, Donaldson, and Komunjer (2011) and Michaels, Rauch, and Redding (2013). Given the unbounded nature of the Fréchet distribution, the production structure does not lead to the full specialization of cities in the production of some sectors, which would make the exposition more involved by inducing censoring at the bottom of the sector-city employment density, without adding substantial insight in the model, given that we do not attempt a structural estimation of the parameters.

\(^{17}\)Note that sector \(j\) spending coming from location \(d\) is equal to the sum of consumer spending \((\alpha_j (1 - \beta) Y_d L_d \pi_{jc})\) and intermediate spending by housing producers \((\alpha_j \beta (1 - \xi) Y_d L_d \pi_{jc})\), so that total spending in \(d\) on \(j\) goods produced in \(c\) is \(\alpha_j (1 - \beta \xi) Y_d L_d \pi_{jc} = \alpha_j w_d L_d \pi_{jc}\).

\(^{18}\)The proportionality coefficients are independent of the sector and city, since \(\theta\) is assumed constant.
is equivalent to the total wage bill in location $c$:

$$w_c L_c = \sum_j \sum_d \pi_{dcj} \alpha_j (1 - \beta \xi) Y_d L_d = \sum_j \alpha_j \pi_{cj} \sum_d w_d L_d = \sum_d w_d L_d \sum_j \alpha_j \pi_{cj} \quad (2.6)$$

Moreover, the housing market must clear in every location, as in Equation (2.2).

**Labor market clearing**  The ratio of labor allocated to sectors $j$ and $j'$ in each city $c$ is given by:

$$\frac{L_{jc}}{L_{j'c}} = \frac{\alpha_j}{\alpha_{j'}} \left( \frac{p_j}{p_{j'}} \right) \theta \frac{D_c}{\theta (\eta_j - \eta_{j'})} \quad (2.7)$$

Total population in a city equals the sum of employment across tradable sectors:

$$\sum_j L_{jc} = L_c \quad (2.8)$$

The labor market clears for the country as a whole:

$$\sum_c L_c = \sum_c \sum_j L_{jc} = \bar{L} \quad (2.9)$$

We can now define the equilibrium of the domestic economy.

**Definition 2.1** (Equilibrium). An equilibrium in the closed economy is defined as an allocation of labor $L_{jc}$ across cities and sectors such that utility is equalized across sites; trading shares satisfy (2.3); labor allocations satisfy (2.7), (2.8) and (2.9); wages satisfy (2.6) and (2.4); tradable prices satisfy (2.5); and housing prices satisfy (2.2).

### 2.2 Implications

#### 2.2.1 Within-Country Specialization

We now investigate the domestic sorting of production generated by the model. Double differencing the spending shares (2.3) from any location $d$ across two goods $j$ and $j'$ and two locations $c$ and $c'$ yields:

$$\left( \frac{\pi_{jc}}{\pi_{j'c}} \right) / \left( \frac{\pi_{jc'}}{\pi_{j'c'}} \right) = \frac{D_c}{D_{c'}} \theta (\eta_j - \eta_{j'}) \quad (2.10)$$

While the absolute unit cost of production is increasing in density $D_c$ due to the need to compensate workers with higher nominal wages, as $D_c$ increases costs increase relatively less fast in sectors with higher $\eta_j$. Denser cities thus have a comparative advantage in sectors that benefit more from agglomeration.\textsuperscript{19} Immediately, this implies:

\textsuperscript{19}Introducing decreasing returns at the establishment level, for example related to the use of a fixed factor in production such as management skill or land, would make these cross-cities, within-country comparative advantage results hold in terms of the number of establishments as well, consistent with our empirical results in section 4.
Lemma 2.1. The share of the labor force employed in higher $\eta_j$ sectors is relatively larger in denser cities:

$$
\left( \frac{L_{jc}}{L_{jc}'} \right) / \left( \frac{L_{j'c}}{L_{j'c}'} \right) = \left( \frac{w_c L_{jc}}{w_{j'c} L_{j'c}'} \right) / \left( \frac{w_{j'c} L_{j'c}}{w_{jc} L_{jc}'} \right) = \left( \frac{\pi_{jc}}{\pi_{j'c}'} \right) / \left( \frac{\pi_{j'c}}{\pi_{jc}'} \right) = \frac{D_c}{D_{c'}}^{\theta(\eta_j - \eta_{j'})}
$$

Equation (2.11) will be key in our empirical estimation of $\eta_j$ for each sector. We use an exogenous shifter of city-level population density to identify the $\eta_j$ for each sector from a city-by-sector level regression of local employment on population density, along with city and sector fixed effects (see Section 3.3).

2.2.2 Cross-Country Specialization

Autarky prices  As a first step toward understanding the relationship between within-country heterogeneity and patterns of trade, we investigate the implications of the model for the country-by-sector level prices in autarky.

Proposition 2.1. The relative price level of two sectors $j$ and $j'$ in the Home country in autarky can be expressed as:

$$
\log \left( \frac{p_j}{p_{j'}} \right) = (\eta_{j'} - \eta_j) \sum_c \omega_{jj',c} \ln(D_c)
$$

where $\omega_{jj',c}$ are bilateral Sato-Vartia weights (Sato, 1976; Vartia, 1976) across any two goods $j$ and $j'$ in city $c$, computed from the export shares:

$$
\omega_{jj',c} = \left( \frac{\pi_{cj} - \pi_{cj'}}{\log(\pi_{cj}) - \log(\pi_{cj'})} \right) / \left( \sum_{d} \frac{\pi_{cj} - \pi_{dj'}}{\log(\pi_{cj}) - \log(\pi_{dj'})} \right)
$$

Proof. See Appendix B.

Conditional on a fixed distribution of city-level densities $D_c$, the closed economy price index in sector $j$ relative to $j'$ is lower when $\eta_j > \eta_{j'}$. Stronger agglomeration forces in a sector increase productivity in all cities, and lower equilibrium prices for any distribution of density. Moreover, we have the following corollary.

Corollary 2.1. Conditional on the vector of $A_c$’s and wages, a more dispersed distribution of $D_c$ across places – defined as second-order stochastic dominance of the density distribution – lowers the price index by more for high $\eta_j$ sectors than for lower $\eta_{j'}$ sectors.

A more dispersed population - i.e. greater variation in $D_c$ – implies relatively more variation in sourcing prices across producing locations for higher $\eta_j$ sectors. Substitution across sourcing cities

\[\text{We can allow for variation in density, conditional on a vector of innate productivity amenities and wages, for example by allowing for an “outside sector” with } \eta_0 = 0 \text{ to be produced with constant productivity } \tilde{A}_0 \text{ in all cities, and thus to determine nominal wages independently of density, with the price of housing adjusting to equalize utilities for dispersed population densities. In that case, with wages pre-determined, a flatter supply curve for housing (as characterized by a lower share of land in production } \xi) \text{ would lead to stronger density variation across cities for a given distribution of wages and innate productivity } \tilde{A}_c.\]

\[\text{This follows immediately from Proposition 2.1, since the log is concave and } \theta > 0. \text{ As in Proposition 1 in Redding and Weinstein (2020), this results from substitutability across suppliers (note we assumed } \theta > \sigma - 1 > 0, \text{ making the price index log sub-modular in } \eta_j \text{ and } D_c'.\]
implies lower relative price indices for more “density-loving” sectors in countries with a more dispersed population. This sub-modularity property of price indices in $\eta_j$ and $D_c$ is at the core of comparative advantage of countries in our global economy.

**Comparative Advantage**  To illustrate the implications of the model for patterns of exports under international trade, we aggregate trade flows at the country level. As in Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016), we study the special case of $N$ countries, indexed by $i$, each composed of a set of regions $c \in C_i$, trading $S$ goods indexed by $j$. We continue to assume that iceberg trade costs are zero across two regions within any country; we also assume trade costs are symmetric and constant across any two regions in two different countries, $\tau_{cc'} = \tau_{ii'} = \tau_{i'i} = \tau_{c'C}$ for $c \in C_i$, $c' \in C_i' \neq i$.

All countries have the same total population $\bar{L} = L_i$ and the same land area $\int_{c \in C_i} B_c = \int_{c \in C_i'} B_c$. We let $B_c = 1$ in each city, so that we simplify the model to the case where $L_c = D_c$. We define $X_{inj}$ as exports from country $i$ to country $n$ in industry $j$, $\bar{w}_{ij} = \frac{\sum_{c \in C_i} w_i L_{jc}}{\sum_{c \in C_i} L_{jc}}$ as the average wage in sector $j$ in country $i$, and $M_i$ as country $i$’s aggregate wage bill, $M_i = \bar{w}_i \bar{L}_i = \sum_j w_{ij} L_{ij}$. We can then state the following aggregation result:

**Proposition 2.2.** Exports of sector $j$ from country $i$ to country $n$ satisfy the following aggregation results

$$X_{inj} = \alpha_j M_n \frac{T_{ij} \bar{w}_{ij}^{-\theta} \tau_{ii}^{-\theta}}{\sum_s T_{sj} \bar{w}_{sj}^{-\theta} \tau_{ns}^{-\theta}}$$

where the country level productivity parameter is:

$$T_{ij} = \left( \sum_{c \in C_i} (A_c D_c^{\eta_j}) \right)^{\frac{\theta}{1 + \theta}} \left( \frac{L_{jc}}{L_{ji}} \right)^{\frac{\theta}{1 + \theta}}$$

Moreover, the aggregate wage bill can be expressed as:

$$M_i = \sum_j w_{ij} L_{ij} = \sum_j \Delta_{ij} \bar{L}_{ij}^{\frac{1}{1 + \theta}} T_{ij}^{\frac{1}{1 + \theta}}$$

where $\Delta_{ij}$, country $i$’s market access in sector $j$, solves the system of $N \times S$ equations:

$$\Delta_{ij} = \left[ \frac{\alpha_j \sum_n M_n \tau_{inj}^{-\theta}}{\sum_s \tau_{is}^{-\theta} \Delta_{js}^{-\theta} L_{sj}^{\frac{1}{1 + \theta}} T_{sj}^{\frac{1}{1 + \theta}}} \right]^{\frac{1}{1 + \theta}}$$

**Proof.** See Appendix B. \(\square\)

The country-sector-level parameter $T_{ij}$ in country $i$ and sector $j$ is endogenous to the distribution of employment shares: Proposition 2.1 and Equation (2.11) immediately imply that $T_{ij}$ is relatively higher for high $\eta_j$ goods in countries with a population more concentrated in a few places, and thus, all else equal, for countries with more variance in sector-neutral productive amenities $A_c$. Even

---

22 This can be seen as a sector-level counterpart to Proposition 1 in Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016).
though we assumed all countries have the same total population, the within-country population distribution drives patterns of cross-country trade.

**Two-Country Case** To build the intuition behind this result, we focus on the the case of two countries, Home and Foreign. First, suppose that Home and Foreign have identical distributions of amenities, $A_c$ and $A^*_c$. Then there will be cross-city trade both within and across countries, but there will be no apparent pattern of inter-industry trade at the country level. More precisely, the distribution of import shares will be the same for any importing destination across all cities in the two countries. Next, assume the distribution of sector-neutral productivity across cities is more even in the Foreign country than at Home. By "more even", we mean that the distribution of Foreign productivity is a "utility-preserving spread," an extension of the "mean-preserving spread" concept defined as:

**Definition 2.2.** $G$ is a "utility-preserving spread" of $G^*$ if in the closed economy, welfare is the same at Home and in Foreign, $\bar{U} = \bar{U}^*$, but the variance of $A_c$ is higher than the variance of $A^*_c$.\(^{23}\)

This implies, from Equation (2.13), that the distribution of population at Home second-order stochastically dominates the distribution in Foreign; the Generalized Lorenz Curve of population in the Foreign economy lies strictly above the Lorenz curve at Home. By Proposition 2.1, the relative prices of higher $\eta_j$ goods are lower in the closed Home economy than in the closed Foreign economy. Equation (2.11) implies that the relative share of employment of high $\eta_j$ sectors is increasing in density, so in the Home country, relatively more workers are active in high $\eta_j$ sectors than in the Foreign country. Aggregating cross-location trade flows to the country level, the Home country will appear to specialize in goods that have a high $\eta_j$'s and import goods with lower $\eta_j$'s. Let the Generalized Lorenz Curve (GLC) of population density be the cumulative distribution function of experienced density, such that $GLC(p)$ is the percentage of the population experiencing a density below the $p$-th percentile of city-level population densities. Then, in particular, in a two-goods setting with identical shares $\alpha_1 = \alpha_2$:

**Corollary 2.2.** Suppose there are two countries, H and F, and two goods $j$ and $j'$ where $\eta_j > \eta_j'$ and $\alpha_j = \alpha_{j'}$. The CDF’s of location-specific amenities in H and F are $G$ and $G^*$. If $G$ is a utility-preserving spread of $G^*$, then the Generalized Lorenz Curve of population-weighted density in H lies strictly below the Generalized Lorenz Curve of population-weighted density in F. Moreover, H is a net exporter of $j$ and F is a net exporter of $j'$.

**2.2.3 From Theory to Measurement: Population-Weighted Density**

One question remains: to link the theory to data, how should we measure the dispersion of $D_c$ at the country-level? The model implies a convenient way to summarize the distribution of city-level densities to a country-level measure. From the equilibrium definition in Section 2.1, the population distribution can be expressed as the labor market clearing (2.9), along with a system of $C$ equations,\(^{23}\)

---

\(^{23}\)One can imagine an experiment with two cities, $c_1$ and $c_2$, where initially $A_{c_2} > A_{c_1}$. Then a utility-preserving spread could involve lowering $A_{c_1}$ by $\epsilon$, and increasing $A_{c_2}$ by $a\epsilon$, where $a$ is chosen so that $\bar{V}_0 = \bar{V}'(\alpha)$. 
one for each city, that depend on the city-level population-weighted densities, the city-level population weighted amenities, and a constant term:\textsuperscript{24}

\begin{equation}
L_c D_c^{\frac{\alpha_j}{\theta_y}} = \sum_j \alpha_j \frac{(A_c D_c^{\eta_j})^{\frac{\alpha_j}{\theta_y}}}{\sum_j (A_c D_c^{\eta_j})^{\frac{\alpha_j}{\theta_y}}} \sum_d L_d D_d^{\frac{\alpha_j}{\theta_y}}
\end{equation}

Since the supply of housing is partially inelastic (due to the use of local land in housing production), the economy has a unique equilibrium when the maximum sector-level density elasticity ($\eta_{\text{max}} = \max_j \eta_j > 0$) is ”not too large” relative to the share of land in housing production ($\xi$); this makes congestion forces strong enough to offset multiple equilibria. When this is the case, as shown in Redding (2016), a location’s density $D_c$ is increasing in its productive amenity $A_c$, since a higher $A_c$ increases the marginal product of labor in any sector, leading to rising nominal wages, population inflows, and land prices, until utility is again equalized. Agglomeration forces, modeled as positive $\eta_j$’s, reinforce this phenomenon, but do not offset it if they are small enough.

Because equilibrium density $D_c$ is increasing in $A_c$, at the country level, a greater dispersion of $A_c$ therefore leads to greater equilibrium $D_c$ dispersion. This is driven by workers relocating from lower to higher $A_c$ and $D_c$ locations. The population density distribution in an economy with more dispersed $A_c$ is second-order stochastically dominated by the population density distribution in an economy with less dispersed $A_c^*$ (see Appendix B).

In the special case where total population is held constant,\textsuperscript{25} and $B_c$ and $A_c$ are uncorrelated, greater dispersion in the exogenous $A_c$’s increases the variance of density, and thus country-level “population-weighted density”:

\begin{equation}
D_i = \int_0^{\max D_c} \frac{L_c^2}{B_c} dH(D_c)
\end{equation}

While, as discussed below, there are several intuitively appealing features of using this as our county-level parameterization of population concentration, it also follows directly from the model. Population-weighted density is thus the country-level observable counterpart of dispersion in $A_c$. This is the measure we estimate next in Section 3, and use in the in the causal analysis in Section 4.1.

3 Data and Descriptive Evidence

3.1 Data Sources

\textbf{Economic Geography in the US} Data on economic activity in the US are collected from the 2016 version of the County Business Patterns (CBP) data set. The CBP contains information on employment, establishment counts, and total payroll in each industry and Core-Based Statistical Area (CBSA). We focus on measures at the NAICS 4-digit level, which are less likely to suffer from sup-

\textsuperscript{24}Using the sectoral price index (2.5) and the spending shares (2.3), immediately yields that the own share of spending in sector $j$ for any location $c \pi_{cj}$ is super-modular in a city’s density and a sector $\eta_j$, since $p_j \propto (\frac{(A_c D_c^{\eta_j})^{\alpha_j}}{\pi_{cj}})^{\frac{1}{1-\theta_y}}$. From the own shares of spending, one can derive an expression for $V$ in equilibrium, as in Redding (2016).

\textsuperscript{25}In the data, we control flexibly for total population.
We compile data on a range of industry-level characteristics from the latest available year in the NBER-CES Manufacturing Industry Database, including capital intensity, the labor share, and average wages. We also include data from the American Community Survey to control for the age and gender breakdown of the workforce as well as detailed measures of the educational attainment of the workforce in each industry. Raster data displaying the distance to bedrock of each 250m grid cell in the US, which we use to construct the instrument for city-level density, are from the International Soil Reference and Information Centre (ISRIC) SoilGrid project.

Trade US State-level international exports from 2016 are collected from the US Census Bureau’s US-ATradeonline database. These data are provided at the NAICS 4-digits level, which is our primary level of analysis across industries. We focus on gross exports flows, as they are the natural counterpart of spending in our theoretical framework. Cross-country trade flows data are obtained from the UN Comtrade Database for all available exporters in 2016, at the HS4 digit level. We map HS4 industries to NAICS-4 industries using the crosswalk developed by Pierce and Schott (2012).

Density Spatial data on global population density are obtained from the LandScan Database. These data are calculated by combining existing demographic and census data with remote sensing imagery, and are released as a raster data set composed of one square-kilometer grid cells. The resultant population count is an ambient or average day/night population count. We use the gridded LandScan data to compute state and country-level estimates of population-weighted density. In our exploration of causality, we also rely on new measures of historical population and city size distributions constructed from data sets recently introduced by Reba, Reitsma, and Seto (2016) and Fang and Jawitz (2018).

---

26 We verify that our results are not sensitive to imputation when using interpolation techniques to impute missing employment data in the CBP.
27 See here: https://www.isric.org/explore/soilgrids.
28 US state-level data were downloaded from the census API at the following link: https://usatrade.census.gov/data/
29 We verify that our results are not sensitive to dropping exporters that are “small countries”, defined as those with population lower than a million.
30 LandScan data can be found here: https://landscan.ornl.gov We use the LandScan data product from 2016.
31 For more information, see here: https://landscan.ornl.gov/documentation. According to LandScan:

ORNL’s LandScan is the community standard for global population distribution. At approximately 1 km resolution (30 x 30 degree), LandScan is the finest resolution global population distribution data available and represents an ambient population (average over 24 hours). [...] The LandScan global population distribution models are a multi-layered, dasymetric, spatial modeling approach that is also referred to as a “smart interpolation” technique. In dasymetric mapping, a source layer is converted to a surface and an ancillary data layer is added to the surface with a weighting scheme applied to cells coinciding with identified or derived density level values in the ancillary data. [...] The modeling process uses sub-national level census counts for each country and primary geospatial input or ancillary datasets, including land cover, roads, slope, urban areas, village locations, and high resolution imagery analysis; all of which are key indicators of population distribution. [...] Within each country, the population distribution model calculates a “likelihood” coefficient for each cell and applies the coefficients to the census counts, which are employed as control totals for appropriate areas. The total population for that area is then allocated to each cell proportionally to the calculated population coefficient.
Additional Data  To include additional controls in our cross-state and cross-country estimates, we compiled US state-level data on educational attainment, age composition, and worker income from the 2016 American Community Survey estimates. At the country level, we also compiled information on educational attainment, urbanization, GDP per capita, and a range of other country-level characteristics from the World Bank’s World Development Indicators and International Monetary Fund’s World Economic Outlook databases, and measures of country-level capital stocks from the Penn World Tables.

3.2 Estimating State and Country Level Density

Using the LandScan data, for both US states and countries, we compute population-weighted density ($D_i$) as:

$$D_i = \sum_{g \in G(i)} \left( \frac{L_g}{\sum_{g' \in G(i)} L_{g'}} \right)$$

where $g$ indexes grid cells and $G(i)$ is the set of grid cells in country (or state) $i$. $L_g$ is the population, according to LandScan, in grid cell $i$. Since all grid cells are the same size, $L_g$ is also the density of grid cell $i$.\footnote{Since grid-cells have an area of one square kilometer (so that the population $L_g$ of a grid-cell is also its density, and $N_{G(i)}$ is also the total area of the country), this is equivalent to computing:}

$$D_i = \text{Population Weighted Density} = \frac{(\text{Mean Density})^2 + (\text{SD of Density})^2}{\text{Mean Density}}$$

This measure is equivalent to weighting the population density of each grid cell in a country or state by its population, and yields a measure of population density that approximates to the expected experienced density of a person in the state or country.\footnote{See Wilson (2012) for a justification of the use of population-weighted density by the United States Census Bureau.}

This is our key state and country-level independent variable of interest. Intuitively, this measure captures the concentration of population within a state or country. For a given total population if people are very concentrated in a few cities this measure will be large whereas if people are is dispersed across many less-dense cities or suburban areas, $D_i$ will be small. Figure 1 plots deciles of $D_i$ for each country around the world. There is substantial variation in $D_i$ across countries, both within continents and within income groups.

3.3 Estimating Sector-Specific Density Affinity

Using industry-by-city level data from the US County Business Patterns (CBP), we estimate the agglomeration elasticity of each tradable manufacturing sector; our estimation follows directly from the model’s Equation 2.11. Because our focus is cross-country trade, and manufactured goods account for the bulk of international exports, we emphasize the existence of substantial within-manufacturing differences in density affinity.

We compute a “density-elasticity” for each industry by estimating the following equation:
where \( c \) indexes cities and \( j \) indexes sectors. \( y_{cj} \) is the (log of the) number of employees, number of establishments, or first quarter aggregate payroll in industry \( j \) and location (city) \( c \). \( \alpha_c \) and \( \gamma_j \) are city and sector fixed-effects, respectively; \( D_c \) is population density at the level of the Core Based Statistical Area (CBSA). \( I_j \) is an indicator that equals one for sector \( j \).

The coefficients of interest are the density elasticities, \( \eta_j \), the key source of industry-level variation in the model (see Equation 2.11). These elasticities capture the extent to which each industry tends to be more or less represented in denser locations.

Since CBSA-level density is likely correlated with a range of other city-level characteristics that might affect industry sorting, it is difficult to interpret the \( \eta_j \)'s at face value. To circumvent this issue, we construct an instrument for CBSA-level density in order to estimate the causal effect of a marginal change in CBSA-level density on industry-specific production. Our instrument is the (log of the) average distance of each CBSA to subterranean bedrock. By exogenously shifting density, we estimate the response of industry specialization to density alone, capturing the causal effect of a marginal change in city-level density on industry-level production.

Lower distance to bedrock in a location eases the land constraint, and can be interpreted as increasing the available share of land \( B_c \) in our theoretical framework; construction often requires a foundation in bedrock and is more difficult when bedrock is deep (Schuberth, 1968; Landau and Condit, 1999). The first stage relationship correlation between CBSA-level density and the log of the distance to bedrock is shown in Figure 2. The correlation coefficient is highly statistically significant (t-statistic = 8.07) suggesting that, consistent with the mechanical impact of distance to bedrock on construction, CBSA-level variation in subterranean bedrock systematically shifts equilibrium population density. The necessary identification assumption is that distance to subterrenian bedrock only affects industry sorting through its impact on ease of construction and hence population density.

Industries with the highest and lowest estimate of density elasticities are listed in Table 1. Since we have six versions of the density elasticities—using employment, establishments, and payroll, estimated using either OLS or IV-2SLS—we report the ten industries with the highest and lowest first principal component of all six elasticity estimates. However, the set of industries is very similar for each elasticity individually. While many of these sectors are intuitive and commonly associated with production in dense cities, in the case of the top sectors, or production away from large cities, in the case of the bottom sectors, they also do not map clearly onto common determinants of comparative advantage. The top of our list features both skill-intensive industries (e.g. Semi-conductor and Other Electronic Component Manufacturing) and industries at the bottom end of the skilled labor

---

34Recent research has suggested the use of underlying geologic characteristics to provide exogenous sources of variation in land supply availability and estimate its economic effects (Rosenthal and Strange, 2008; Saiz, 2010; Duranton and Turner, 2018) However, most of the existing research has focused on within-city variation in geological features to instrument for urban shape, rather than cross-metropolitan areas variation.

35While this assumption seems likely, we also verify that the results are similar after controlling for other ground and soil characteristics (e.g. characteristics of soil content, agricultural suitability, etc.). These estimates and their possible parameterizations are available upon request.
Figure 2: **Distance to Bedrock and Population Density.** The figure is a binned scatter plot. It reports the correlation between log of distance to bedrock and log of population density at the CBSA level. The t-statistic is 8.07.

requirement distribution (e.g. Bakeries and Tortilla Manufacturing). The same is true for capital intensity.\textsuperscript{36} Indeed, the first principal component of our elasticities is not significantly correlated with capital intensity, several measures of skill intensity, or the age or gender breakdown of employment. This suggests that our analysis is not just capturing well-understood determinants of comparative advantage, something that we verify in more detail empirically in the next sections.

Figure 3 shows the distribution of establishments in the top and bottom ten sectors listed in Table 1 across the US. For each CBSA $c$ and sector $j$, we compute:

$$
\text{Representation}_{cj} = \left( \frac{\sum_{j \in T,B} \text{Establishments}_{cj}}{\sum_{j} \text{Establishments}_{cj}} \right) / \left( \frac{\sum_{c} \sum_{j \in T,B} \text{Establishments}_{cj}}{\sum_{c} \sum_{j} \text{Establishments}_{cj}} \right)
$$

where $T$ and $B$ are the set of ten highest and lowest $\eta_j$ sectors respectively. This normalization captures the over- or under-representation of top or bottom sectors in city $c$ by normalizing the share of city $c$ manufacturing establishments that belong to $j \in T/B$ by the overall share of manufacturing establishments that belong to $j \in T/B$ in the US. Otherwise, larger sectors would appear highly represented everywhere, and all sectors would appear well represented in larger cities.

Figure 3a shows the geographic distribution of low-$\eta_j$ sectors; they are disproportionately located in Upper Midwest and Central and Northern Plains regions (purple-shaded regions). High-$\eta_j$ sectors, displayed in Figure 3b, are disproportionately located on the East and West coasts, as well as in

\textsuperscript{36}Moreover, motor vehicle manufacturing and navigational equipment manufacturing are both at the top of Nunn (2007)'s list of contract intensive industries, but are at opposite ends of our list.
Table 1: The Ten Most and Least Density Elastic Industries

<table>
<thead>
<tr>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>First PC, $\eta_j$</td>
<td>NAICS Code</td>
<td>Industry Name</td>
<td>First PC, $\eta_j$</td>
<td>NAICS Code</td>
<td>Industry Name</td>
</tr>
<tr>
<td>3.077139</td>
<td>3342</td>
<td>Communications Equipment Manufacturing</td>
<td>-1.719853</td>
<td>3325</td>
<td>Hardware Manufacturing</td>
</tr>
<tr>
<td>2.931949</td>
<td>3344</td>
<td>Semiconductor and Other Electronic Component Manufacturing</td>
<td>-1.941343</td>
<td>3161</td>
<td>Leather and Hide Tanning and Finishing</td>
</tr>
<tr>
<td>2.815478</td>
<td>3345</td>
<td>Navigational, Measuring, Electromedical, and Control Instruments Manufacturing</td>
<td>-2.293531</td>
<td>3114</td>
<td>Specialty Food Manufacturing</td>
</tr>
<tr>
<td>2.754001</td>
<td>3327</td>
<td>Screw, Nut, and Bolt Manufacturing</td>
<td>-2.347334</td>
<td>3331</td>
<td>Machinery Manufacturing</td>
</tr>
<tr>
<td>2.673068</td>
<td>3222</td>
<td>Converted Paper Product Manufacturing</td>
<td>-2.607069</td>
<td>3361</td>
<td>Motor Vehicle Manufacturing</td>
</tr>
<tr>
<td>2.618597</td>
<td>3231</td>
<td>Printing and Related Support Activities</td>
<td>-2.749151</td>
<td>3112</td>
<td>Grain and Oilseed Milling</td>
</tr>
<tr>
<td>2.464373</td>
<td>3118</td>
<td>Bakeries and Tortilla Manufacturing</td>
<td>-2.819922</td>
<td>3274</td>
<td>Lime and Gypsum Product Manufacturing</td>
</tr>
<tr>
<td>2.258956</td>
<td>3219</td>
<td>Other Wood Product Manufacturing</td>
<td>-3.192425</td>
<td>3111</td>
<td>Animal Food Manufacturing</td>
</tr>
<tr>
<td>2.190134</td>
<td>3121</td>
<td>Beverage Manufacturing</td>
<td>-4.159338</td>
<td>3122</td>
<td>Tobacco Manufacturing</td>
</tr>
</tbody>
</table>

Notes: The density elasticity measure is the first principal component from our six elasticity estimates for each sector.

cities in Texas and parts of the Midwest. There is significant variation within regions and states as well. Indeed, almost all states have locations in which both high and low $\eta_j$ sectors are disproportionately produced.

Discussion Our baseline empirical results do not take a strong stance on industry-specific characteristics driving variation in the “density affinity” of manufacturing industries. Existing estimates of the impact of urban agglomeration on productivity find substantial variation across sectors, and our results corroborate this (Shefer, 1973; Nakamura, 1985; Rosenthal and Strange, 2004; Faggio, Silva, and Strange, 2017). Recent work has proposed industry-level variables that determine the extent to which sectors benefit from agglomeration and production in denser cities; these include education and skill requirements (Davis and Dingel, 2014) or capital intensity (Gaubert, 2018). An important distinction between most recent work and our estimates is that we restrict attention to tradable manufacturing sectors; therefore, the fact that high-skilled services, for example, are disproportionately located in cities is outside the scope of our analysis.

An alternative source of density affinity is the intensive use of differentiated local services (Abdel-Rahman and Fujita, 1990; Abdel-Rahman and Fujita, 1993; Abdel-Rahman, 1994; Abdel-Rahman, 1996). In this framework, density facilitates the production of non-tradable services (e.g. (Clark, 1945), and hence service-reliant sectors sort into dense cities. Anecdotally, many large companies
Figure 3: **Representation of Low- and High- \( \eta_j \) Sectors Across US Cities.** Both (a) and (b) are US CBSA-level maps. (a) displays the relative representation of low-\( \eta_j \) sectors, the ten sectors with the lowest first principal component of our six density elasticity estimates. (b) displays the relative representation of high-\( \eta_j \) sectors, the ten sectors with the lowest first principal component of our six density elasticity estimates. These sectors are listed in Table 1.

justify their relocation in large cities by their readily available diversity of services producers.\(^{37}\) Some manufacturing sectors may locate in dense cities because of their improved ability to efficiently source from non-tradable sectors in the larger local market.

Another potential determinant of variation in density affinity is the extent to which each sector relies on raw materials (e.g. minerals, agriculture) as inputs. Sectors that rely on immobile natural resources might be less able to locate in cities and reap the benefits of agglomeration (Ades and Glaeser, 1995); locating in urban centers for these sectors would mean paying high transportation costs on inputs. Finally, dense cities might be particularly productive places for innovation and R&D (e.g. Duranton and Puga, 2004). If this is the case, the density affinity measure might be capturing the extent to which each sector benefits from innovation and the role of R&D in the production process. We investigate the potential contribution of these mechanisms in Section 4.5.

4 Empirical results: Density and Trade

4.1 Estimation Framework

We now examine the impact of within-country population distribution on patterns of trade. We investigate whether population-weighted density, \( D_{ij} \), is a systematic source of comparative advantage.

\(^{37}\)See Bruce Nollop, *Wall Street Journal - The Experts*, April 25, 2016: “As companies focus on their core competencies, they can benefit greatly from cities’ networks of service providers.”
The primary estimating equation is:

\[
\text{Exports}_{ij} = \alpha_i + \gamma_j + \beta \cdot \eta_{IV}^j \cdot \ln(D_i) + X'_{ij}\Gamma + e_{ij} \tag{4.1}
\]

where \(i\) indexes states or countries and \(j\) indexes sectors. The unit of observation is a country (or state) by sector pair. The dependent variable is total exports in sector \(j\) from state or country \(i\). The independent variable of interest is an interaction term between (i) IV estimates of sector-level density affinity (\(\eta_{IV}^j\)) and (ii) log of state or country-level population weighted density (\(\ln(D_i)\)). The density affinity of all NAICS-4 sectors were estimated using Equation (3.1) and the instrumental variables strategy outlined in Section 3.3. All specifications include sector and state or country fixed effects; we also include a range of controls that vary at the state-by-sector or country-by-sector level (\(X'_{ij}\)) chosen to absorb potential omitted variables.

The coefficient of interest is \(\beta\) in all estimates of (4.1). If \(\beta > 0\), it implies that countries with greater population-weighted density have a revealed comparative advantage in “density-loving” sectors. This framework follows the “regression-based index” of comparative advantage summarized in French (2017), as used, among others, by Nunn (2007) or Bombardini, Gallipoli, and Pupato (2012). Following Silva and Tenreyro (2006), we use the Poisson pseudo-maximum likelihood (PPML) estimator as our baseline specification, but show throughout that results are similar using OLS.\(^{38}\)

Finally, in Section 4.4 we propose an instrumental variables strategy that exploits variation in historical population and city size distributions as shifters of modern population density.

### 4.2 US State-Level Estimates

The over-representation of some manufacturing sectors in dense areas in the United States might stem from either local supply or local demand conditions. Our hypothesis focuses on the supply side, by suggesting that denser cities are relatively more efficient in the production of “density-loving” industries. If this is the case, dense areas within the US should not only attract relatively more employment and production in these industries, but also export significantly more of them internationally.

Even across US states, there is substantial variation in the pattern of population distribution. Moreover, while many models of international trade consider the entire US as a single “point” or observation, different parts of the US specialize in vastly different industries (see e.g. Irwin (2017) for a long-term perspective). Thus, as a first test of our hypothesis that regions with greater population-weighted density specialize in the export of density-loving industries, we estimate Equation (4.1) at the US state level.\(^{39}\)

Table 2 presents estimates of Equation (4.1) at the state-by-sector level.\(^{40}\) Panel A reports Poisson

---

\(^{38}\)As shown by Fally (2015), the Poisson pseudo-maximum likelihood estimation method has the additional benefit of ensuring that predicted trade flows satisfy the “adding up” constraint implicit in gravity models of trade.

\(^{39}\)While some recent studies have attempted to estimate export data at the metropolitan level (see e.g. the database constructed by Tomer and Kane (2014)), most trade flows data are still collected at a broader level of aggregation. The smallest level of consistent and exhaustive trade reporting in the United States is the state.

\(^{40}\)We restrict the sample to manufacturing industries only because of the far better coverage of the County Business Patterns in these sectors, in order to focus on tradable goods, and to avoid results being driven by agricultural exports.
### Table 2: State-Level Trade, Baseline Estimates

<table>
<thead>
<tr>
<th>Strategy for estimation of density affinity:</th>
<th>(\eta_j) computed using industry-level employee</th>
<th>(\eta_j) computed using industry-level number of establishments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Column 1</td>
<td>(D_i \times \eta_j)</td>
<td>0.612***</td>
</tr>
<tr>
<td>(0.145)</td>
<td></td>
<td>0.539***</td>
</tr>
<tr>
<td>Column 2</td>
<td>(D_i \times \eta_j)</td>
<td>0.146*</td>
</tr>
<tr>
<td>(0.0734)</td>
<td></td>
<td>0.129*</td>
</tr>
<tr>
<td>Column 3</td>
<td>(D_i \times \eta_j)</td>
<td>0.142*</td>
</tr>
<tr>
<td>(0.0725)</td>
<td></td>
<td>0.120*</td>
</tr>
<tr>
<td>Column 4</td>
<td>(D_i \times \eta_j)</td>
<td>0.124</td>
</tr>
<tr>
<td>(0.0738)</td>
<td></td>
<td>0.864**</td>
</tr>
<tr>
<td>Column 5</td>
<td>(D_i \times \eta_j)</td>
<td>0.839**</td>
</tr>
<tr>
<td>(0.0685)</td>
<td></td>
<td>0.358</td>
</tr>
<tr>
<td>Column 6</td>
<td>(D_i \times \eta_j)</td>
<td>0.349</td>
</tr>
<tr>
<td>(0.0793)</td>
<td></td>
<td>0.363</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.756</td>
<td>0.758</td>
</tr>
<tr>
<td>Factor Intensity Controls</td>
<td>No No</td>
<td>0.757</td>
</tr>
<tr>
<td>State Level Controls</td>
<td>Yes No</td>
<td>0.758</td>
</tr>
<tr>
<td>Industry Level Controls</td>
<td>Yes No</td>
<td>0.757</td>
</tr>
<tr>
<td>State FE</td>
<td>Yes Yes</td>
<td>0.760</td>
</tr>
<tr>
<td>Industry FE</td>
<td>Yes Yes</td>
<td>0.760</td>
</tr>
<tr>
<td>States</td>
<td>50 50</td>
<td>50 50</td>
</tr>
<tr>
<td>Observations</td>
<td>4,182 4,132</td>
<td>4,182 4,132</td>
</tr>
</tbody>
</table>

Notes: The unit of observation is a state-by-sector pair. The coefficient of interest is the coefficient on an interaction between state-level population weighted density and sector-level density affinity computed using the bedrock IV and city-level employment in columns 1-5 and establishments in columns 6-7. Panel A reports Poisson pseudo-maximum likelihood estimates while Panel B reports OLS estimates. All specifications include state and sector fixed effects, along with other controls listed at the bottom of each column. Standard errors, clustered at the state level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.

Maximum likelihood estimates while Panel B reports OLS estimates with log of exports as the outcome variable. Across specifications, we find that the coefficient of interest is positive and statistically significant, suggesting that US states with greater population-weighted density have a comparative advantage in density-loving industries.

Column 1 presents the coefficient of interest when only \(\eta_j^{IV} \times \ln(D_i)\)—the interaction between state-level population weighted density and industry-level density affinity—is included on the right hand side (along with state and industry fixed effects). The remaining specifications investigate the robustness of this baseline result to the inclusion of additional controls.

In order to address the concern that the results are driven by state-level differences in education and comparative advantage in high-skill industries, in column 2 we include a series of interactions between state-level educational attainment and sector-level skill demand. In particular, we separately interact the share of people in each state who have achieved a (i) high school degree, (ii) a bachelor’s degree, and (iii) a graduate degree, with the share of people employed in each sector (i) that have a high school degree or (ii) that have at least a college degree. The inclusion of these six interactions has little effect on our coefficient of interest.

In column 3, we control for a series of state-level variables interacted \(\eta_j^{IV}\) in order to investigate from low-density states.
whether the baseline result is driven by some omitted state-level characteristic. These controls include (log of) the median household income; (log of) state-level population; the share of inhabitants with high school, bachelor, and graduate degree; and the share of young people, aged 18-30. It is possible, for example, that denser states are also just wealthier and that this drives the correlation in column 1. However, the coefficient of interest remains similar after including these state-level controls.

In order to address the potential for omitted industry-level characteristics, in column 4 we control for a series of industry-level characteristics interacted with \( \ln(D_i) \). These covariates, computed for each manufacturing industry in the US, are the value of installed capital per worker, (log of) the average employee compensation, the share of workers with at least a college degree, the average age of employees, and the gender breakdown of employment.

In column 5 we include all 17 controls mentioned thus far and again, the coefficient of interest remains very similar. It does, however, lose statistical significance in Panel B when we use an OLS regression model and log of exports as the outcome variable; this is driven by a larger standard error rather than a decline in coefficient magnitude.

In columns 6-7 we repeat the specifications from columns 1 and 5—the specifications without any controls and the specification with all controls—and construct the “density affinity” measure using industry-level establishment data rather than employment data. The number of establishments is a potentially less noisy measure of industry-level production across space than employment, and moreover is never suppressed in the CBP data. Reassuringly, in both columns 6 and 7 and in both Panels A and B, our coefficient of interest is positive and highly significant.

This first set of results demonstrates that US states that exhibit a more spatially concentrated population export relatively more in sectors whose production is concentrated in denser metropolitan areas. According to our estimates, a one-standard deviation increase in the density interaction in the fully controlled specification increases the dependent variable by 0.139 standard deviations when computed using the elasticity with respect to employment and 0.295 when computed using the elasticity with respect to establishments. These results suggest that the distribution of population may affect not only the economic geography of a region, but also its international industrial specialization.

Sensitivity and Robustness In Table A1 we further test the sensitivity of the baseline result by estimating a series of additional specifications. Each coefficient reported in Table A1 is the result from a separate regression. First, we consider the robustness of the result to alternative measures of \( \eta_j \). While for our baseline results, we primarily focused on a version of \( \eta_j^{IV} \) computed from city-level employment data, we also compute \( \eta_j^{IV} \) using city-level establishment and payroll data. Rows 1-3 of Table A1 reproduce the baseline results using versions of \( \eta_j^{IV} \) computed using employment, establishments, or payroll data; the coefficient of interest is qualitatively very similar across specifications. In column 2, we add the full set of controls from Table 2 to the right side of the regression and again the coefficients of interest are similar.

Yet another way we can compute the independent variable of interest is to calculate the \( \eta_j \) using OLS, rather than IV, estimates of Equation (3.1). While we prefer our IV strategy because there are
likely many city-level characteristics that are correlated with density and might influence the sorting of production, for clarity we report the full set of specifications from rows 1-3 using \( \eta_j \)'s computed using OLS instead of IV-2SLS. These specifications tell a very similar story. All 12 coefficients presented in Table A1 are statistically significant at below the 1% level.

Finally, while in the baseline results, we exclude state-industry pairs with zero exports, the estimates are very similar if we include the zeroes. These results are reported in Table A2 and are reassuring since they imply that are results are not driven by the omission of observations with no trade.

4.3 Country-Level Estimates

We now turn to the main results of the paper: the relationship between density and patterns of cross-country trade. Estimates of (4.1) in which the units of observation are country-industry pairs are reported in Table 3. Panel A presents Poisson maximum likelihood estimates while Panel B reports estimates from an OLS model. The coefficient of interest in a specification without controls is presented in column 1; it is positive and highly significant. Countries with a more concentrated population distribution—the darker countries in Figure 1—have a revealed comparative advantage in density-loving sectors.

Columns 2-6 investigate the robustness of the result to the inclusion of a series of controls in order address potential concerns due to omitted variable bias. In column 2, we control for traditional determinants of comparative advantage, including capital and skill intensity (Romalis, 2004). Since data on the country-level capital stock is only available for 90 countries, the sample size of the regression is reduced; nevertheless, the coefficient of interest is almost exactly identical.

In column 3 we control for a series of country-level characteristics interacted with the sector-level density elasticity measure, \( \eta_j^{IV} \). These are included to account for the fact that population-weighted density is potentially related to other country-level characteristics that may affect comparative advantage. In particular, we control for (the log of) country-level total population, educational attainment, urbanization, the share of population employed in agriculture, the share of population employed in service production, (log of) per capita GDP (PPP adjusted), and a rule of law index, all interacted with \( \eta_j^{IV} \). Again, the coefficient of interest is very similar after the inclusion of these controls and remains highly statistically significant.

Next, we investigate the robustness of the result to the inclusion of sector-level controls. Analogous to the concern that there is an omitted country-level variable that might bias the result, we might be concerned that our the \( \eta_j^{IV} \) capture some industry-level characteristic that is omitted from the specification in column 1. To address this, we control for the same industry-level controls as in Table 2, interacted with country-level measures of population-weighted density \( D_i \). Reassuringly, the

\[\text{follow the Penn World Tables with an industry's average level of capital intensity obtained from the NBER-CES Manufacturing database. We also interact measures of educational attainment at the country level with our estimates of the skill intensity of an industry in US data computed from the share of high school and college attainment of workers in the industry in the American Community Survey data. The coefficient of interest is also similar if only individual country-level controls or smaller sets of country-level controls are included on the right-hand side, but to conserve space we do not report these specifications.}\]
Table 3: Country-Level Trade, Baseline Estimates

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variable is Total Exports from the Country-Sector</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Panel A: Outcome Variable is Total Exports (Thousands), PML Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D_i \times \eta_j )</td>
<td>0.456***</td>
<td>0.464***</td>
<td>0.757***</td>
<td>0.462***</td>
<td>0.765***</td>
</tr>
<tr>
<td></td>
<td>(0.111)</td>
<td>(0.110)</td>
<td>(0.0849)</td>
<td>(0.0710)</td>
<td>(0.0731)</td>
</tr>
<tr>
<td><strong>Panel B: Outcome Variable is log(Exports), OLS Model</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( D_i \times \eta_j )</td>
<td>0.104**</td>
<td>0.105**</td>
<td>0.288***</td>
<td>0.122***</td>
<td>0.262***</td>
</tr>
<tr>
<td></td>
<td>(0.0487)</td>
<td>(0.0524)</td>
<td>(0.0645)</td>
<td>(0.0454)</td>
<td>(0.0627)</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.814</td>
<td>0.796</td>
<td>0.793</td>
<td>0.816</td>
<td>0.797</td>
</tr>
<tr>
<td>Factor Intensity Controls</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Country Level Controls</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry Level Controls</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Country FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Countries</td>
<td>134</td>
<td>90</td>
<td>107</td>
<td>134</td>
<td>83</td>
</tr>
<tr>
<td>Observations</td>
<td>10,464</td>
<td>7,241</td>
<td>8,542</td>
<td>10,332</td>
<td>6,674</td>
</tr>
</tbody>
</table>

Notes: The unit of observation is a country-by-sector pair. The coefficient of interest is the coefficient on an interaction between country-level population weighted density and sector-level density affinity computed using the bedrock IV and city-level employment. Panel A reports Poisson pseudo-maximum likelihood estimates while Panel B reports OLS estimates. All specifications include country and sector fixed effects, along with other controls listed at the bottom of each column. Standard errors clustered at the country level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.

Coefficient of interest is again very similar after the inclusion of these controls.

In column 5, we include all controls mentioned thus far on the right-hand side of the regression. Due to missing covariates, the sample size is reduced to 83 countries, yet the coefficient of interest remains positive and highly significant.

This first set of country-level results suggests that the distribution of population within countries is a potentially important determinant of comparative advantage and patterns of trade. Our estimates from column 2, when only factor endowment controls are included, imply that a one standard deviation increase in the density interaction increases the outcome variable by 0.113 standard deviations. This is slightly larger in magnitude than the coefficient on the capital interaction, which implies a standardized beta coefficient of 0.109.\[^{43}\] In the specification with all controls included, the coefficient of interest increases and implies a beta coefficient on the density interaction of 0.276.

**Sensitivity and Robustness** We present a series of robustness tests to assess the sensitivity of the baseline estimates. First, in Table A3 we reproduce all baseline results after including continent-by-continental controls.

\[^{43}\]Interestingly, our estimates of the magnitudes of comparative advantage due to factor endowments is very similar to Nunn (2007), who estimates a beta coefficient on an analogous capital interaction of 0.105.
industry fixed effects. This specification compares the comparative advantage of countries within the same continent that have different population distributions; it was motivated by the fact that, visually, there is some variation across continents in population-weighted density (see Figure 1). The results remain very similar and, if anything, are more precisely estimated after the inclusion of these additional fixed effects.

We next report a series of specification checks that are analogous to the state-level tests reported in Table A1; these are displayed in Table A4. We report estimates from versions of the baseline regression equation that use all possible strategies to compute sector-level density affinity. The only difference from the state-level robustness tests is that we add additional columns, columns 3-4, in which we exclude countries at the bottom end of the population and income distribution. The lowest income countries likely also have lower quality data and the smallest or poorest countries might have extreme values of either density or trade values. The country-level results are very similar across specifications, suggesting that the baseline results are not driven by the details of our variable. Of the 24 specifications presented in Table A4, the coefficient of interest loses statistical significance in only one case (Row 4, Column 2).

Finally, the results are very similar if we include country-industry pairs with zero exports. These results are reported in columns 3-4 of Table A2 and are very similar to our baseline estimates.

### 4.4 Endogeneity

This section proposes an instrument for population-weighted density and reports instrumental variable estimates of our baseline specification. The goal of introducing an instrument is to make sure that the baseline results are not driven by reverse causality. That is, it is possible that the composition of a state or country’s exports has feedback effects and shapes its economic geography; we would then find a positive coefficient on our density interaction, but it would be incorrect to interpret the relationship as evidence that density is a source of comparative advantage. To rule out the possibility that our results capture the effect of trade on economic geography, we use characteristics of a state or country’s historical population distribution to construct instruments for the population distribution today. While characteristics of a country’s historical population distribution predict its modern population distribution, it seems unlikely that modern patterns of trade, which developed largely after World War II, had a direct effect on the population distribution in 1900 (e.g. Irwin, 2017).

The ideal instrument for our purposes would be a historical measure of population weighted density, analogous to our contemporary measure. We construct such a measure for each US state using estimates of the historical US population distribution presented in Fang and Jawitz (2018). Fang and Jawitz (2018) combine historical census data with population modeling techniques to construct a spatially explicit distribution of the US population for each decade since 1790. Using this gridded...
data set, we compute the population weighted density of each US state in 1900 ($D_{i1900}$). The first stage estimating equation is thus:

$$\left(\eta_{i}^{IV} \cdot \ln(D_i)\right) = \xi \cdot \eta_{i}^{IV} \cdot \ln(D_{i1900}) + \alpha_i + \gamma_j + + X_{ij}' \Gamma + \epsilon_{ij}$$

(4.2)

where we hypothesize $\xi > 0$ if the historical state-level population distribution is a strong predictor of the modern population distribution.

Our state-level IV-2SLS estimate of Equation (4.1), where the first stage estimating equation is (4.2), is presented in column 1 of Table 4. The coefficient estimate is positive, statistically significant, and similar in magnitude to the OLS estimates, suggesting that our state-level findings are not driven by reverse causality. Moreover, the first stage relationship is also strong; the Kleibergen-Paap first stage F-statistic is 25.159.

While it is possible to estimate the historical population weighted density of each US state, to our knowledge this is not possible at the country level. Therefore, in order to adapt the logic of our identification strategy to the country-level analysis, we also introduce a second set of instruments. We determined the location and population of cities around the world in 1900 using historical data collected by Chandler (1987), and recently digitized by Reba, Reitsma, and Seto (2016), in order construct an intuitively similar measure with global coverage from the city-level data. High $D_i$ corresponds to having a high city population concentrated in a relatively small number of cities. For each state, we therefore compute the total population across all cities in the Chandler (1987) data ($p_{i1900}$), as well as the inverse number of cities ($c_{i1900}$). We include both, as well as their interaction ($p_{i1900} \cdot c_{i1900}$), interacted with $\eta_j$, as excluded instruments. We expect $p_{i1900} \cdot c_{i1900} \cdot \eta_j$ to be positively correlated with $D_i \cdot \eta_j$, the endogenous variable, since a high value of $p_{i1900} \cdot c_{i1900}$ implies that in 1900 the state had high overall city population concentrated in a small number of cities. This is the analog to population weighted density we compute from the city-level data.

The first stage estimating equation using the city-level data is:

$$\left(\eta_{i}^{IV} \cdot \ln(D_i)\right) = \zeta \cdot \eta_{i}^{IV} \cdot p_{i1900} + \xi \cdot p_{i1900} \cdot \eta_{j}^{IV} + \phi \cdot p_{i1900} \cdot c_{i1900} \cdot \eta_{j}^{IV} + \alpha_i + \gamma_j + + X_{ij}' \Gamma + \epsilon_{ij}$$

(4.3)

and we hypothesize that $\phi > 0$. States (and below, countries) with a high historical urban population concentrated in a small number of cities should—if the logic of the instrument is correct—have higher population-weighted density today.

State-level IV-2SLS estimates of Equation (4.1) with this second instrumentation strategy are reported in columns 2-3 of Table 4. The sample is reduced to 39 states because 11 states have no cities in the Chandler (1987) data in 1900. Nevertheless, the estimates remain positive and highly significant. Since $p_{i1900}$ (total urban population in 1900), one of the excluded instruments, will likely be mechanically correlated with modern population, we control for modern (log of) country population

---

46 We select the year 1900 for comparability with our country-level IV estimates, which have additional data constraints and are reported below.

47 1900 was chosen because it is the oldest year with broad and global coverage.
Table 4: State-Level Trade, IV Estimates

<table>
<thead>
<tr>
<th>Strategy for estimation of density affinity:</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dc x ηj</td>
<td>0.231**</td>
<td>0.149**</td>
<td>0.288***</td>
<td>1.098***</td>
<td>0.657*</td>
<td>0.951***</td>
</tr>
<tr>
<td>ln(population) x ηj</td>
<td>-0.106</td>
<td>-0.0921</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

K-P F-Statistic

| State FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Industry FE | Yes | Yes | Yes | Yes | Yes | Yes |
| States   | 48  | 39  | 39  | 48  | 39  | 39  |
| Observations | 4,182 | 4,132 | 4,182 | 4,132 | 4,132 | 4,182 |

Notes: The unit of observation is a state-by-sector pair. The coefficient of interest is the coefficient on an interaction between state-level population weighted density and sector-level density affinity computed using the bedrock IV and city-level employment in columns 1-3 and establishments in columns 4-6. All estimates report IV-2SLS estimates. In columns 1 and 3, the excluded instrument is an interaction between sector-level density affinity and state-level population weighted density computed from the US 1900 population distribution. In columns 2-3 and 5-6, the excluded instruments are the total urban population in the state in 1900, the inverse number of cities, and the interaction between the two. The Kleibergen-Paap F-statistic for each first stage regression is reported at the bottom of each column. Standard errors, clustered at the state level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.

Interacted with ηj in column 3; the coefficient of interest remains positive and significant. Finally, in columns 4-6, we repeat the results from columns 1-3 except in all cases use the version of ηjIV estimated from data on establishments rather than data on employment; the results are qualitatively identical.

Next, we turn to IV-2SLS estimates of our country-level results. Across countries, we rely exclusively on the instruments constructed from the Chandler (1987) city-level data. However, it is worth noting that across US states, our instrument constructed from the Chandler (1987) data and our estimate of historical population weighted density are highly positively correlated; the binned partial correlation plot is reported as Figure 4. Country-level IV-2SLS estimates of Equation (4.1) are presented in Panel A of Table 5; the first stage estimating equation is Equation 4.3 and first stage estimates are reported in Panels B. For comparison, Panel C reports OLS estimates.

Our baseline country-level IV-2SLS estimate is reported in column 1 of Table 5. The coefficient estimate is positive and significant, supporting the argument that density is a source of comparative advantage and that our baseline estimates are not driven by reverse causality. Reassuringly, and following the state-level analysis, in the first stage specification we find that φ > 0 while the direct effects of p1900i and c1900i are both negative. The IV estimate, however, is larger in magnitude than the
### Table 5: Country-Level Trade, IV Estimates

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variable is Total Exports from the State-Sector</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Full Sample</td>
<td>Excluding Bottom 10% by Population</td>
<td>Excluding Bottom 10% by Income</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Panel A: IV-2SLS Estimates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_i \times \eta_j$</td>
<td>0.517**</td>
<td>0.279**</td>
<td>0.411***</td>
<td>0.319***</td>
<td>0.404**</td>
<td>0.214**</td>
</tr>
<tr>
<td></td>
<td>(0.236)</td>
<td>(0.117)</td>
<td>(0.196)</td>
<td>(0.116)</td>
<td>(0.185)</td>
<td>(0.0894)</td>
</tr>
<tr>
<td>$\ln(\text{population}) \times \eta_j$</td>
<td>-0.0895**</td>
<td>-0.0434</td>
<td>-0.0887**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0366)</td>
<td>(0.0407)</td>
<td>(0.0346)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>Panel B: First Stage Estimates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(p_{i, 1900} \times (c_{i, 1900}) \times \eta_j$</td>
<td>0.787**</td>
<td>1.021***</td>
<td>0.797**</td>
<td>1.091***</td>
<td>1.119***</td>
<td>1.153***</td>
</tr>
<tr>
<td></td>
<td>(0.344)</td>
<td>(0.312)</td>
<td>(0.338)</td>
<td>(0.345)</td>
<td>(0.382)</td>
<td>(0.356)</td>
</tr>
<tr>
<td>$p_{i, 1900} \times \eta_j$</td>
<td>-0.614***</td>
<td>-0.728***</td>
<td>-0.634***</td>
<td>-0.766***</td>
<td>-0.705***</td>
<td>-0.787***</td>
</tr>
<tr>
<td></td>
<td>(0.213)</td>
<td>(0.180)</td>
<td>(0.211)</td>
<td>(0.189)</td>
<td>(0.227)</td>
<td>(0.192)</td>
</tr>
<tr>
<td>$c_{i, 1900} \times \eta_j$</td>
<td>-8.705**</td>
<td>-10.54***</td>
<td>-8.782**</td>
<td>-11.36***</td>
<td>-12.43***</td>
<td>-11.83***</td>
</tr>
<tr>
<td></td>
<td>(3.868)</td>
<td>(3.497)</td>
<td>(3.807)</td>
<td>(3.868)</td>
<td>(4.304)</td>
<td>(3.998)</td>
</tr>
<tr>
<td>R-Squared</td>
<td>0.095</td>
<td>0.463</td>
<td>0.115</td>
<td>0.474</td>
<td>0.127</td>
<td>0.527</td>
</tr>
<tr>
<td></td>
<td><strong>Panel C: OLS Estimates</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_i \times \eta_j$</td>
<td>0.134**</td>
<td>0.196***</td>
<td>0.169***</td>
<td>0.181***</td>
<td>0.129**</td>
<td>0.198***</td>
</tr>
<tr>
<td></td>
<td>(0.0624)</td>
<td>(0.0709)</td>
<td>(0.0608)</td>
<td>(0.0676)</td>
<td>(0.0635)</td>
<td>(0.0719)</td>
</tr>
<tr>
<td>$\ln(\text{population}) \times \eta_j$</td>
<td>-0.0753**</td>
<td>-0.0175</td>
<td>-0.0861**</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0334)</td>
<td>(0.0344)</td>
<td>(0.0332)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Countries</td>
<td>86</td>
<td>86</td>
<td>77</td>
<td>77</td>
<td>78</td>
<td>78</td>
</tr>
<tr>
<td>Observations</td>
<td>7022</td>
<td>7022</td>
<td>6281</td>
<td>6281</td>
<td>6379</td>
<td>6379</td>
</tr>
</tbody>
</table>

Notes: The unit of observation is a country-by-year pair. Panel A reports IV-2SLS estimates, Panel B reports first stage estimates, and Panel C reports OLS estimates. The coefficient of interest is the coefficient on an interaction between country-level population weighted density and sector-level density affinity computed using the bedrock IV and city-level employment. $p$ is the log of the total urban population in 1900 and $c$ is the inverse number of cities. All specifications include country and sector fixed effects, along with other controls listed at the bottom of each column. Sample restrictions are noted in the column header. The Kleibergen-Paap F-statistic for each first stage regression is reported at the bottom of Panel B. Standard errors clustered at the country level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.

An OLS estimate. One explanation for this is that the IV estimate is capturing a particular local average treatment effect. For example, it could be the case that countries whose modern economic geography is highly correlated with economic geography in 1900 are also countries that industrialized early, and are very specialized in industries that fit their population distribution. This would generate IV estimates that are larger than OLS, since the the instrument captures variation across countries whose specialization is very responsive to their population distribution.

Another possible explanation, as noted above, is that variation in the instruments is correlated with the error term in the second stage regression. Indeed, the instruments are constructed from historical population data and likely capture variation in total population and not only variation in
Following the control strategy in our baseline results, in column 2 we include an interaction term between the (log of) present day population and \( \eta IV_j \) as a control. The IV coefficient is smaller in magnitude in column 2 and more precisely estimated. While it remains larger than the OLS estimate, it is no longer statistically distinguishable.

A potential concern with using the Chandler (1987) data is that data quality and coverage are likely different for different sets of countries. In particular, it is likely of lower quality for smaller and lower income countries, which might be more likely to have cities excluded from the data. To make sure this is not driving the result, in columns 3-4 and 5-6 we repeat the specifications from columns 1-2 after dropping countries in the bottom 10% of the population and income distribution respectively. Reassuringly, our estimates remain very similar. The results are also similar if we drop countries in the bottom 20 or 25% of the distribution (not reported).

While it seems unlikely that urban geography in 1900 was caused by modern patterns of trade, it is nevertheless a possibility that historical urban geography affected modern determinants of comparative advantage other than \( D_i \). If elements of countries’ historical economic geography influence modern determinants of comparative advantage though channels other than features of modern economic geography, it would be a violation of the exclusion restriction. Nevertheless, the robustness of our result to the battery of controls and specifications in the previous section, as well as the broadly similar results using these historical instruments, indicates that density is a potentially important causal determinant of patterns of trade.

4.5 Mechanisms: What drives density affinity?

We next turn to potential mechanisms underpinning the baseline results. While in the main specification we relied on a reduced-form measure of industry-level “density affinity,” in this section we explore which industry characteristics potentially underlie the baseline estimates. First, some recent work has highlighted the greater skill and level of human capital in cities (Davis and Dingel, 2014). It is worth noting that in the baseline specification, we were careful to control flexibly for the potential role of variation in skill or education, both across sectors and across countries. In column 1 of Table 6, we report the coefficient on the interaction between population-weighted density and the share of employment in each industry in the US with a college degree. The coefficient on this interaction is statistically insignificant; we also find no evidence that education is driving the result if we break the industry-level education measure into a larger number of discrete bins (not reported). Thus, we do not find strong evidence that our estimate of comparative advantage are driven by cross-industry heterogeneity in education.

Another potential determinant of our density affinity measure is the extent to which each sector relies on differentiated local services. Population density might facilitate the productive provision of services and sectors that rely more on local services may therefore benefit disproportionately from density (Abdel-Rahman and Fujita, 1990; Abdel-Rahman and Fujita, 1993; Abdel-Rahman, 1994; Abdel-Rahman, 1996). Our estimates lend some support to this hypothesis. Within the United States, we find that manufacturing industries in which services comprise a large share of total intermediate
inputs tend to locate in denser areas.\textsuperscript{48} When we turn to the trade data, however, service reliance does not explain the export patterns of high-$\eta_j$ sectors (column 2). The coefficient on the interaction between population-weighted density and industry-level service intensity is in fact negative and far from statistically significant.

Certain industries may locate away from dense cities if they rely on immobile natural resources (e.g. Ades and Glaeser, 1995). These sectors might be less able to benefit from urban externalities and variation in natural resource dependence across industries might drive our variation in density affinity. Anecdotally, the sectors at the bottom of our “density affinity” list seem to be those that source extensively from natural resources (see Table 1). To investigate this, we compute the share of natural resource inputs for each manufacturing sector using the US input-output tables. The coefficient on the interaction term between population-weighted density and industry-level natural resource dependence is negative and significant (column 3 of Table 6), suggesting that indeed denser countries export less in sectors that rely on natural resources. This is consistent with the idea that resource-reliant sectors optimally locate away from urban centers and that dense countries hence are

\textsuperscript{48}We compute each sector’s non-tradable input share from the Bureau of Labor Statistics input-output tables.
disproportionately productive in industries that do not rely on natural resources.

A final potential mechanism is the role of research and development (R&D) in production. Industries rely differentially on R&D expenditure and innovation in the production process. If cities facilitate innovation (e.g. Duranton and Puga, 2001; Duranton and Puga, 2004), then sectors that rely disproportionately on R&D might be especially productive in dense cities. Therefore, our baseline estimates might be capturing the role of density in facilitating R&D. To investigate this, for each sector we compile data on (i) R&D spending per worker and (ii) the share of employees in science, technology, engineering, and mathematical (STEM) fields from the Brookings Advanced Industries database.49 Again, we include an interaction term between both measures and country-level density in our baseline country-level estimating equation; the estimates are reported in column 4 of Table 6. Both interactions are positive and statistically significant, suggesting that density may play a role in facilitating R&D and that denser places specialize in the export of R&D intensive sectors.

Does the combination of these channels explain our baseline estimates? In columns 5-6, we include all variables from columns 1-4 of Table 6 on the right hand side of the regression, along with the industry-level reduced form density affinity interacted with population-weighted density. If our proposed mechanisms fully explained the baseline results, we would expect the coefficient on the density affinity variable to be zero. However, it remains positive and statistically significant, whether density-affinity is measured using US employment data (column 5) or data on establishments (column 6). Thus, we find suggestive evidence that (i) the role of cities in facilitating R&D and (ii) heterogeneity in industry-specific natural resource dependence are important channels; however, they do not fully explain our baseline results, suggesting that additional and un-observed industry characteristics are also at play. Uncovering industry-level characteristics that drive sorting with respect to density strikes us as a potentially interesting area for additional exploration.

5 Conclusion

This paper argues that some countries “specialize in density”: in the language of the factor endowment literature, countries that exhibit an abundance of dense cities export relatively more in “dense-city-intensive sectors.” Most analysis of sources of comparative advantage in international trade have emphasized aggregate variation in country-level endowments or production technologies. Our theory and empirical results, however, suggest that even when two countries look identical in the aggregate, they may still specialize in vastly different industries because the domestic distribution of factors of production is a key determinant of comparative advantage.

Within the United States, we uncover substantial heterogeneity in the density-affinity of tradable sectors. While some sectors are disproportionately located in large cities, others are more frequently found in small cities or suburban areas. We find that US States and countries with higher population-weighted density – that is, with a more concentrated population – export relatively more in sectors with high density affinity. The results are robust to the inclusion of a battery of controls and are

similar using an instrumental variables strategy that exploits variation in countries’ historical city size distribution.

While the impact of trade on sub-national regions is increasingly well understood, the impact of economic geography and domestic heterogeneity on patterns of trade has been less well documented. The goal of this study is to explore one facet of this relationship, and to show that domestic economic geography can itself be a key driver of international trade flows.

References


Pierce, Justin R and Peter K Schott (2012). “A concordance between ten-digit US Harmonized System Codes and SIC/NAICS product classes and industries”. In: Journal of Economic and Social Measurement 37.1, 2, pp. 61–96.
Schuberth, Christopher J (1968). The geology of New York City and environs. Published for the American Museum of Natural History [by] Natural History Press.
### Table A1: State-Level Trade, Alternative Specifications

<table>
<thead>
<tr>
<th></th>
<th>Column (1)</th>
<th>Column (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
<td>Total Exports (Thousands)</td>
<td>Total Exports (Thousands)</td>
</tr>
<tr>
<td><strong>η, computed using:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employment, IV</td>
<td>0.612***</td>
<td>0.538***</td>
</tr>
<tr>
<td></td>
<td>(0.145)</td>
<td>(0.199)</td>
</tr>
<tr>
<td>Establishments, IV</td>
<td>3.508***</td>
<td>3.241***</td>
</tr>
<tr>
<td></td>
<td>(0.541)</td>
<td>(0.660)</td>
</tr>
<tr>
<td>Payroll, IV</td>
<td>0.335***</td>
<td>0.295***</td>
</tr>
<tr>
<td></td>
<td>(0.0753)</td>
<td>(0.111)</td>
</tr>
<tr>
<td>Employment, OLS</td>
<td>0.788***</td>
<td>0.459***</td>
</tr>
<tr>
<td></td>
<td>(0.236)</td>
<td>(0.172)</td>
</tr>
<tr>
<td>Establishments, OLS</td>
<td>2.650***</td>
<td>1.766***</td>
</tr>
<tr>
<td></td>
<td>(0.462)</td>
<td>(0.401)</td>
</tr>
<tr>
<td>Payroll, OLS</td>
<td>0.504***</td>
<td>0.307***</td>
</tr>
<tr>
<td></td>
<td>(0.169)</td>
<td>(0.117)</td>
</tr>
</tbody>
</table>

**All Controls**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>State FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>4,182</td>
<td>4,132</td>
</tr>
</tbody>
</table>

**Notes:** The unit of observation is a state-by-sector pair. Each coefficient is an estimate from a separate regression. The coefficient of interest is the coefficient on an interaction between state-level population weighted density and sector-level density affinity using the strategy listed on the left side of the table. All reported specifications are Poisson pseudo-maximum likelihood estimates and include state and sector fixed effects, along with other controls listed at the bottom of each column. Standard errors clustered at the state level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.
Table A2: Baseline Results Including Observations with No Exports

<table>
<thead>
<tr>
<th></th>
<th>US State-Level</th>
<th>Country-Level</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Outcome Variable:</strong></td>
<td>Exports</td>
<td>Exports</td>
</tr>
<tr>
<td><strong>Model:</strong></td>
<td>PML</td>
<td>OLS</td>
</tr>
<tr>
<td><strong>D_j x η</strong></td>
<td>0.612***</td>
<td>0.425**</td>
</tr>
<tr>
<td></td>
<td>(0.145)</td>
<td>(0.169)</td>
</tr>
<tr>
<td><strong>State FE</strong></td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Country FE</strong></td>
<td>-</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Industry FE</strong></td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Observations</strong></td>
<td>4,250</td>
<td>11,122</td>
</tr>
<tr>
<td><strong>R-squared</strong></td>
<td>0.709</td>
<td>0.823</td>
</tr>
</tbody>
</table>

Notes: The unit of observation is a state-industry pair (columns 1-2) or a country-industry pair (columns 3-4). The coefficient of interest is the coefficient on an interaction between state- or country-level population weighted density and sector-level density affinity computed using the bedrock IV and city-level employment. In columns 1 and 3, the outcome variable is total exports and in columns 2 and 4, it is the inverse hyperbolic sine of total exports. Observations with zero exports are included in the estimation. Standard errors clustered at the state (columns 1-2) or country (columns 3-4) level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.
Table A3: Country-Level Trade, Including Continent $\times$ Industry Fixed Effects

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><strong>Dependent Variable is Total Exports from the Country-Sector</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$D_i \times \eta_j$</td>
<td>Panel A: Outcome Variable is Total Exports (Thousands), PML Model</td>
<td>Panel B: Outcome Variable is log(Exports), OLS Model</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.191)</td>
<td>(0.181)</td>
<td>(0.163)</td>
<td>(0.0986)</td>
<td>(0.158)</td>
</tr>
<tr>
<td></td>
<td>0.412**</td>
<td>0.403**</td>
<td>0.486***</td>
<td>0.380***</td>
<td>0.491***</td>
</tr>
<tr>
<td></td>
<td>(0.0667)</td>
<td>(0.0770)</td>
<td>(0.0757)</td>
<td>(0.0627)</td>
<td>(0.0826)</td>
</tr>
<tr>
<td>Factor Intensity Controls</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Country Level Controls</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry Level Controls</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Country FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Industry x Continent FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Countries</td>
<td>134</td>
<td>90</td>
<td>107</td>
<td>134</td>
<td>83</td>
</tr>
<tr>
<td>Observations</td>
<td>10,464</td>
<td>7,159</td>
<td>8,542</td>
<td>10,332</td>
<td>6,674</td>
</tr>
</tbody>
</table>

Notes: The unit of observation is a country-by-sector pair. The coefficient of interest is the coefficient on an interaction between country-level population weighted density and sector-level density affinity computed using the bedrock IV and city-level employment. Panel A reports Poisson pseudo-maximum likelihood estimates while Panel B reports OLS estimates. All specifications include country and continent-by-sector fixed effects, along with other controls listed at the bottom of each column. Standard errors clustered at the country level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.
Table A4: Country-Level Trade, Alternative Specifications

<table>
<thead>
<tr>
<th>Sample:</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable is Total Exports (Thousands)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Full Sample</td>
<td>Excluding countries with</td>
<td>Excluding bottom 10%</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>pop &lt; 1 million</td>
<td>income</td>
<td></td>
</tr>
<tr>
<td>( \eta_j ) computed using:</td>
<td>Employment, IV</td>
<td>0.456***</td>
<td>0.774***</td>
<td>0.457***</td>
</tr>
<tr>
<td></td>
<td>(0.111)</td>
<td>(0.0720)</td>
<td>(0.111)</td>
<td>(0.111)</td>
</tr>
<tr>
<td></td>
<td>Establishments, IV</td>
<td>1.594***</td>
<td>1.836***</td>
<td>1.594***</td>
</tr>
<tr>
<td></td>
<td>(0.361)</td>
<td>(0.262)</td>
<td>(0.362)</td>
<td>(0.361)</td>
</tr>
<tr>
<td></td>
<td>Payroll, IV</td>
<td>0.248***</td>
<td>0.401***</td>
<td>0.248***</td>
</tr>
<tr>
<td></td>
<td>(0.0640)</td>
<td>(0.0408)</td>
<td>(0.0640)</td>
<td>(0.0640)</td>
</tr>
<tr>
<td></td>
<td>Employment, OLS</td>
<td>0.292**</td>
<td>0.135</td>
<td>0.292**</td>
</tr>
<tr>
<td></td>
<td>(0.147)</td>
<td>(0.0881)</td>
<td>(0.147)</td>
<td>(0.147)</td>
</tr>
<tr>
<td></td>
<td>Establishments, OLS</td>
<td>0.793**</td>
<td>0.480**</td>
<td>0.792**</td>
</tr>
<tr>
<td></td>
<td>(0.329)</td>
<td>(0.225)</td>
<td>(0.329)</td>
<td>(0.328)</td>
</tr>
<tr>
<td></td>
<td>Payroll, OLS</td>
<td>0.224**</td>
<td>0.105*</td>
<td>0.224**</td>
</tr>
<tr>
<td></td>
<td>(0.0985)</td>
<td>(0.0580)</td>
<td>(0.0987)</td>
<td>(0.0984)</td>
</tr>
</tbody>
</table>

All Controls: No Yes No No
Country FE: Yes Yes Yes Yes
Industry FE: Yes Yes Yes Yes
Observations: 10,464 6,674 9,277 9,515

Notes: All reported coefficients are from regressions at the country-by-sector level. Each coefficient is an estimate from a separate regression. The coefficient of interest is the coefficient on an interaction between country-level population weighted density and sector-level density affinity computed using the strategy listed on the left hand side of each row. All reported specifications are Poisson pseudo-maximum likelihood estimates and include country and sector fixed effects, along with other controls listed at the bottom of each column. Sample restrictions are noted in the column header. Standard errors clustered at the country level, are reported in parentheses. *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.
Figure 4: **Correlation Between Both US State-Level Instruments.** This figure presents the partial correlation, conditional on state and industry fixed effects, between (i) log of US state-level population weighted density in 1900, estimated from the Fang and Jawitz (2018) data set, and (ii) the interaction between total 1900 city population and the inverse number of cities, estimated from the Chandler (1987) data set.
B Derivations and proofs

B.1 Housing market

Out of nominal disposable income $Y_c$, a worker in city $c$ spends a constant share $p_{hc}Y_c = \beta Y_c$ on the non-tradable good produced in city $c$, and a constant share $(1 - \beta)Y_c = X_c$ on the basket of tradable sectors, with sub-shares $a_i X_c = p_{jc} c_j^\iota$ on each sector $j$. Each landowner faces a price $p_{hc}$ for housing and a cost of $P$ for the numeraire input. Each landowner then uses an amount $X_{hc} = \gamma (1 - \zeta) (\frac{p_{hc}}{P})^{\frac{1}{\gamma}}$ of tradable inputs, and aggregate housing supply is: $H^s(c) = \frac{B_c (\frac{p_{hc}}{P})^{\frac{1}{\gamma}}}{\xi}$. Equalizing supply and demand yields equilibrium housing prices in each city (equation 2.2):

$$p_{hc}^{\frac{1}{\gamma}} = \frac{\beta L_c Y_c}{B_c \xi^\gamma - \xi}$$

Landowners in a city receive proceeds from real estate sales $\beta Y_c L_c$, out of which they spend $P X_{hc} = (1 - \zeta) \beta Y_c L_c$ on the final good, while accruing rents $r_c B_c = \xi \beta Y_c L_c$. $r_c$ is defined as the Ricardian rent per unit of land, increasing in local population density and local disposable income. Using the spatial equilibrium condition and the fact that all land rents are fully rebated to local workers, we have:

$$Y_c = \bar{U} P^{1 - \beta} p_{hc}^{\beta} = \bar{U} P^{1 - \beta} \left( \frac{\beta L_c Y_c}{B_c \xi^\gamma - \xi} \right)^{\beta \xi} = \bar{U} P^{1 - \beta} \left( \frac{\beta L_c Y_c}{B_c \xi^\gamma - \xi} \right)^{\beta \xi}$$

and thus

$$w_c = P (1 - \beta \xi) \bar{U} P^{1 - \beta} \left( \frac{\beta L_c Y_c}{B_c \xi^\gamma - \xi} \right)^{\beta \xi} \propto P \times D_c^{\beta \xi} \xi^\gamma$$

B.2 Comparative advantage of cities

Cost minimization by consumers in any location $d$ implies, in the absence of trade costs and using standard Eaton-Kortum algebra (Costinot, Donaldson, and Komunjer, 2011; Michaels, Rauch, and Redding, 2013):

$$p_{dj}(\omega) = \min \{ p_{dcj}(j); c \in C \}$$

The probability that the unit cost is less than $p$ for variety $\omega$ of good $j$ produced in $c$ is:

$$F_{jc}(p) = \mathbb{P}(\frac{w_c}{\zeta} < p) = 1 - e^{-\left( \frac{w_c}{\zeta A_c D_{ij}^{\theta}} \right)^{\theta}}$$

The probability that the minimal cost for variety $\omega$ of good $j$ is less than $p$ is thus:

$$F_{j}(p) = 1 - (\Pi_{c \in C} (1 - F_{jc}(p))) = 1 - e^{-\sum_{c'} (A_{c} D_{ij}^{\theta}) w_{c'}^{\theta} p^{\theta}}$$

and the probability that location $c$ is the lowest cost supplier for variety $\omega$ for location $d$ is:

$$\mathbb{P}(\frac{w_c}{\zeta_{jc}} \leq \min \{ p_{dcj}(j); c \in C \}) = \frac{A_{c} D_{ij}^{\theta} w_{c}^{\theta}}{\sum_{c'} (A_{c} D_{ij}^{\theta}) w_{c'}^{\theta}}$$
From the Fréchet distribution assumption and the Constant Elasticity of Substitution structure on demand allocation within good \( j \), standard algebra then implies that the share of spending on varieties from location \( c \) in sector \( j \) must be equal across all locations \( d \):

\[
\pi_{dcj} = \pi_{cj} = \frac{p_{cj}X_{dcj}}{X_{dj}} = \frac{(A_cD_c^\eta_j)^\theta w_c^{-\theta}}{\sum_c(A_cD_c^\eta_j)^\theta w_c^{-\theta}} \tag{B.1}
\]

where \( \pi_{dcj} \) denotes spending in city \( d \) on goods in sector \( j \) produced in city \( c \), equation 2.3 in the model.

B.3 Proposition 2.1

The derivation borrows from the definition of the unified price index in Redding and Weinstein (2020). Using spending shares 2.3, and the definition of the price index 2.5, we obtain:

\[
\frac{\pi_{cj}}{\pi_{cj'}} = \left( \frac{P_j}{P_j'} \right)^\theta \frac{(A_cD_c^\eta_j)^\theta w_c^{-\theta}}{(A_cD_c^\eta_j')^\theta w_c'^{-\theta}}
\]

Re-expressing and taking logs on both sides:

\[
\frac{\log(\pi_j)}{\log(\pi_{j})} - (\eta_j' - \eta_j) \log(D_c) = \frac{1}{\theta}
\]

Multiplying both sides by \( \pi_{cj} - \pi_{cj'} \), and using that in autarky \( \sum_c \pi_{cj} = 1 \), summing over all cities \( c \) and rearranging yields the Sato-Vartia relative price:

\[
\sum_{c \in C} \left( \frac{\pi_{cj} - \pi_{cj'}}{\log(\pi_{cj}) - \log(\pi_{cj'})} \right) \log\left( \frac{P_j}{P_j'} \right) = (\eta_j' - \eta_j) \sum_{c \in C} \left( \frac{\pi_{cj} - \pi_{cj'}}{\log(\pi_{cj}) - \log(\pi_{cj'})} \right) D_c
\]

and, rearranging, we obtain the “Sato-Vartia” relative price expression in proposition 2.1.

B.4 Population density dispersion

Because equilibrium density \( D_c \) is increasing in \( A_c \), at the country level, greater dispersion of \( A_c \) therefore leads to greater equilibrium \( D_c \) dispersion, as workers reallocate from lower to higher-\( A_c \), higher-\( D_c \) locations. The population density distribution in an economy with more dispersed \( A_c \) is second-order stochastically dominated by the population density distribution in an economy with less dispersed \( A_c^* \) (see B).

Formally, suppose there are two countries, \( H \) and \( F \), and define \( H(d) \) as the share of the total

---

\(^{50}\)Given the unbounded nature of the Fréchet distribution, the production structure does not lead to the full specialization of cities in the production of some sectors, which would make the exposition more involved by inducing censoring at the bottom of the sector-city employment density, without adding substantial insight in the model, given that we do not attempt a structural estimation of the parameters.
population living in cities with density below \( d \) in \( H \), the high-amenity-dispersion economy:

\[
H(d) = \frac{\sum_{c \in C} L_c \mathbb{1}(\frac{L_c}{B_c} \leq d)}{L}
\]

Let \( H^*(d) \) be its counterpart in \( F \). Then, for any \( d \), we have:

\[
\int_0^d H(s) ds \geq \int_0^d H^*(s) ds
\]

For any percentile \( p \), there is a corresponding density threshold \( H^{-1}(p) = d \). Let the Generalized Lorenz Curve (GLC) of population density be the function:

\[
GLC(p) = \int_0^p H^{-1}(q) dq, \text{ for } p \in [0, 1]
\]

Integration by parts yields:

\[
GLC(p) \leq GLC^*(p) \forall p
\]

The GLC of density in a country with a higher dispersion of population lies strictly below that of a country with a more concentrated distribution of population. Note that we have, by a change of variable:

\[
GLC(p) = \frac{\sum_{c \in C} \frac{(L_c)^2}{B_c} \mathbb{1}(H(\frac{L_c}{B_c}) \leq p)}{L}
\]

### B.5 Proposition 2.2

We assume, as in Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016), that iceberg trade costs are nil within a country, and symmetric (at the country-level) across any two locations in two different countries. The proof follows the structure of Ramondo, Rodríguez-Clare, and Saborío-Rodríguez (2016), extended to a case with many sectors.

We obtain a natural extension of equation 2.4 in a world of many countries, namely that for any city \( c \) in country \( i \), the wage bill in sector \( j \) satisfies:

\[
w_c L_{jc} = \alpha_j \sum_n \sum_{c' \in C_s} \frac{(A_{c'D_{c}^j})^\theta w_{c'}^\theta \tau_{in}^{-\theta}}{\sum_{c' \in C_s} (A_{c'D_{c}^j})^\theta w_{c'}^\theta \tau_{sn}^{-\theta}} \sum_{d \in C_n} w_d L_d
\]

We rewrite equation (B.2) as:

\[
w_c = \left( \frac{A_{c'D_{c}^j})}{L_{jc}} \right)^{\frac{1}{\theta}} \Delta_{ij}
\]

where \( \Delta_{ij} \) is a country-sector level variable indexing market access in sector \( j \) and country \( i \):

\[
\Delta_{ij}^{1+\theta} = \alpha_j \sum_n \sum_{c' \in C_s} \frac{\tau_{in}^{-\theta}}{\sum_{c' \in C_s} (A_{c'D_{c}^j})^\theta w_{c'}^\theta \tau_{sn}^{-\theta}} \sum_{d \in C_n} w_d L_d
\]
We can use the fact that:

\[
\sum_{d \in C_n} w_d L_d = \sum_{d \in C_n} \sum_{k} w_d L_{dk}
\]

and equation (B.2) to re-express \( \Delta_{ij} \):

\[
\Delta_{ij}^{1+\theta} = \alpha_j \sum_{n} \frac{\tau_{in}^{-\theta} \sum_{d \in C_n} \sum_{k} L_{kd} \left( \frac{A_{ij} L_{ik}^{\delta}}{L_{dk}} \right)^{1+\theta} \Delta_{nk}}{\sum_{n} \Delta_{nk} L_{nk}^{1+\theta} \sum_{d \in C_n} \sum_{k} \left( A_{ij} D_{ik}^{\delta_j} \right)^{1+\theta} \left( \frac{L_{kd}}{L_{nk}} \right)^{1+\theta}} \tag{B.5}
\]

where \( L_{nk} = \sum_{d \in C_n} L_{dk} \). We define the following objects, that depend on the equilibrium distribution of population within a country:

\[
T_{ij} = \left( \sum_{c \in C_i} \left( A_{ij} D_{ij}^{\delta_j} \right)^{1+\theta} \left( \frac{L_{ij}}{L_{ji}} \right)^{1+\theta} \right)^{1+\theta} \tag{B.6}
\]

\[
M_i = \sum_j \Delta_{ij} L_{ij}^{1+\theta} T_{ij}^{1+\theta} \tag{B.7}
\]

Note then that we can re-express equation (B.5) as a system of equations in \( M_n, T_{sj}, L_{sj}, \) and \( \Delta_{sj} \):

\[
\Delta_{ij}^{1+\theta} = \alpha_j \frac{\sum_{n} M_n \tau_{in}^{-\theta}}{\sum_{s} \tau_{is}^{-\theta} \sum_{j} \Delta_{ij} T_{ij}^{1+\theta} L_{ij}^{1+\theta}} \tag{B.8}
\]

We make note that \( M_i \) corresponds to the total tradable wage bill in a country:

\[
\sum_{c \in C_i} w_c L_c = \sum_{c \in C_i} \sum_{j} w_c L_{cj} = \sum_{j} \Delta_{ij} L_{ij}^{1+\theta} T_{ij}^{1+\theta} = M_i \tag{B.9}
\]

We now use fact (B.9) to derive the bilateral export flows from country \( i \) to country \( n \) in sector \( j \), by using the fact that exports of good \( j \) from any city \( c \in C_i \) to any city \( d \in C_n \) are given by:

\[
x_{cdj} = \alpha_j w_d L_d \frac{(A_{ij} D_{ij}^{\delta_j})^{1+\theta} w_c^{-\theta} \tau_{in}^{-\theta}}{\sum_{s} \tau_{is}^{-\theta} \sum_{j} \left( A_{ij} D_{ij}^{\delta_j} \right)^{1+\theta} w_c^{-\theta}}
\]

Summing over cities, using (B.5), (B.7) and (B.6), yields, after rearranging:

\[
X_{inj} = \sum_{c \in C_i} \sum_{d \in C_n} x_{cdj} = \alpha_j M_n \tau_{in}^{-\theta} \frac{\Delta_{ij}^{1+\theta} T_{ij}^{1+\theta} L_{ij}^{1+\theta}}{\sum_{s} \Delta_{sj}^{-\theta} T_{sj}^{1+\theta} L_{sj}^{1+\theta}} \tag{B.10}
\]
We next derive the average wage in country $i$ and sector $j$:

$$w_{ij} = \frac{\sum_{c \in C_i} w_c L_{cj}}{\sum_{c \in C_i} L_{cj}}$$

by using equation (B.2), again summing over all cities in country $i$ and using the same manipulations:

$$w_{ij} = \sum_n X_{ijn} \frac{\sum_{c \in C_i} X_{ijn} L_{ij}}{L_{ij}} = \alpha_j \frac{\sum_n M_n T_{ijn}^{-\theta} \Delta_{ij}^{-\theta} T_{ij}^{\frac{1}{\theta}} L_{ij}^{-\frac{1}{\theta}}}{\sum_s \Delta_s^{-\theta} T_{sj}^{\frac{1}{\theta}} L_{sj}^{\frac{1}{\theta}}}$$ (B.11)

and, using the system (B.8) and substituting, we obtain:

$$w_{ij} = \Delta_{ij} \left( \frac{T_{ij}}{L_{ij}} \right)^{\frac{1}{1+\theta}}$$ (B.12)

Plugging (B.12) into equation (B.10) yields proposition 2.2.