

NBER WORKING PAPER SERIES

CITIES, HETEROGENEOUS FIRMS, AND TRADE

Jan David Bakker
Alvaro Garcia Marin
Andrei V. Potlogea
Nico Voigtländer
Yang Yang

Working Paper 32377
<http://www.nber.org/papers/w32377>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2024

This paper combines independent work from Garcia-Marin, Potlogea, Voigtländer, and Yang (2020) and Bakker (2020). We would like to thank Dominick Bartelme, Fabian Eckert, Pablo Fajgelbaum, Oleg Itskhoki, Tim Kehoe, Fernando Parro, Thomas Sampson, Oren Ziv as well as numerous seminar and conference audiences. Bakker thanks his advisors and members of his committee, Kalina Manova, Ferdinand Rauch, Daniel Sturm, Gabriel Ulyssea and Tony Venables. Francisco Diaz-Valdes and Alexander Valenzuela provided superb research assistance. We gratefully acknowledge research support from UK Research and Innovation (ESRC New Investigator Grant, Grant Reference ES/W008017/1, project title "Cities, Trade and Productivity"), from the UCLA Ziman Center's Rosalinde, Arthur Gilbert Program in Real Estate, Finance and Urban Economics, and from CONICYT-Chile through project Fondecyt Regular 1191536. Corresponding author: Nico Voigtländer (nico.v@ucla.edu). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2024 by Jan David Bakker, Alvaro Garcia Marin, Andrei V. Potlogea, Nico Voigtländer, and Yang Yang. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Cities, Heterogeneous Firms, and Trade

Jan David Bakker, Alvaro Garcia Marin, Andrei V. Potlogea, Nico Voigtländer, and Yang Yang

NBER Working Paper No. 32377

April 2024

JEL No. F1,R1

ABSTRACT

Does international trade affect the growth of cities, and vice versa? Assembling disaggregate data for four countries, we document a novel stylized fact: Export activity is disproportionately concentrated in larger cities – even more so than overall economic activity. We rationalize this fact by marrying a standard quantitative spatial economics model with a heterogeneous firm model that features selection into the domestic and the export market. Our model delivers novel predictions for the bi-directional interactions between trade and urban dynamics: On the one hand, trade liberalization shifts employment towards larger cities, and on the other hand, liberalizing land use raises exports. We structurally estimate the model using data for the universe of Chinese manufacturing and French firms. We find that trade policies have quantitatively meaningful impacts on urban outcomes and vice versa, and that the aggregate effects of trade and urban policies differ from more standard models that do not account for the interaction between trade and cities. In addition, a distinguishing prediction of our model – which we confirm in the data – is that local trade elasticities vary systematically with city size, so that a country's aggregate trade elasticity depends on the spatial distribution of production within its borders.

Jan David Bakker
Bocconi University
Department of Economics
Via Roentgen 1
20136 Milan
Italy
jan.bakker@unibocconi.it

Alvaro Garcia Marin
Universidad de los Andes
Monseñor Álvaro del Portillo 12455
Las Condes, Santiago
Chile
alvarof.garciam@gmail.com

Andrei V. Potlogea
School of Economics
University of Edinburgh
Room 2.08, 31 Buccleuch Place
Edinburgh, EH8 9JT
Scotland
Andrei.Potlogea@ed.ac.uk

Nico Voigtländer
Anderson School of Management
University of California, Los Angeles
110 Westwood Plaza
C513 Entrepreneurs Hall
Los Angeles, CA 90095
and NBER
nico.v@anderson.ucla.edu

Yang Yang
International Monetary Fund
700 19th St NW
Washington, DC 20431
yyang5@imf.org

1 Introduction

Over the last decades, two mega-trends have shaped economies across the globe: rapid urbanization and a surge in international trade.¹ These trends have traditionally been examined by two separate strands of literature – international trade and economic geography.² We ask whether the growth of cities and trade are related and what implications such a connection would have for welfare and income inequality.

We examine the relationship between city size and ‘export intensity’ (city-level exports relative to overall city-level revenues). Using data for China, France, Brazil, and the United States, we establish a novel stylized fact: Larger cities systematically export a higher fraction of their output, even after controlling for geographic characteristics and market access (Figure 1). Moreover, the majority of this association can be attributed to variation *within* industries. Thus, differences in industrial composition (such as car manufacturing vs. real estate services) cannot explain why exporting is concentrated in larger cities. Within industries, in turn, the intensive margin (i.e., export intensity of exporting firms) is relatively unimportant. Thus, systematic differences in variable trade costs cannot account for our stylized fact. Instead, it is driven by the extensive margin: In larger cities, a higher proportion of firms participate in export markets. This implies that heterogeneous firms play an important role in giving rise to the novel stylized fact.

Motivated by our empirical findings, we propose a simple model that marries core mechanisms from international trade and urban economics. The model features multiple locations and – in each location – heterogeneous firms that select into serving the domestic and export market (Melitz, 2003). We first highlight this mechanism in a simple framework without domestic trade costs (Rosen-Roback) and symmetric countries. A straightforward prediction is that in equilibrium, firms in larger cities are more productive. Crucially, however, differences in *average* productivity across cities are not sufficient to produce differences in export *intensity*. This may seem surprising; it results because average productivity affects the domestic and the export entry thresholds proportionately so that exports relative to total output of a location remains the same. However, adding a well-known empirical regularity to the model can replicate the disproportionate concentration of exporters in larger cities: We introduce differences in the upper tail of the productivity distribution across locations, reflecting the empirical pattern documented in the literature on city size and productivity (see Combes, Duranton, Gobillon, Puga, and Roux, 2012) – a productivity pattern that we also document in our data (see Figure 2). Thus, in our model, entrants in more

¹The average urbanization rate in the world grew from 37 to 57 percent between 1970 and 2022. During the same period, exports as a share of GDP almost tripled, from 13 to 31 percent (<https://data.worldbank.org/indicator>).

²More recently, a literature at the intersection of these fields has emerged, but it focuses more broadly on the spatial effects of trade (and its sectoral patterns) within countries, thus speaking only indirectly to urbanization. We discuss this literature in detail below.

productive (and hence larger) cities draw from Pareto distributions with thicker upper tails.³ This increases the proportion of high-productivity firms – that select into exporting – in larger cities, leading to a disproportionate concentration of exporting in larger cities.⁴

Our simple model delivers another novel prediction: The (international) trade elasticity varies across space and is *smaller* for larger cities. This is a surprising result; the spatial heterogeneity in trade elasticities is driven by the different shapes of the local productivity distributions.⁵ While our key assumptions naturally lead to this prediction, it does not follow from workhorse spatial models without differences in the upper tail of the firm productivity distribution (e.g., Allen and Arkolakis, 2014; Redding, 2016). Thus, we can use this prediction to validate the core assumption of our model vis-à-vis alternative mechanisms. We find strong support, using the identification strategy from Pierce and Schott (2016) for Chinese exporters. Our results show the trade elasticity is about 25% smaller in large cities (90th percentile of population) relative to small cities (10th percentile).

In addition to explaining these patterns in the data, our model delivers novel predictions that are relevant for policy-making, highlighting important interactions between trade and economic geography. First, a reduction in international trade costs shifts economic activity towards larger cities. Second, the model leads to a novel mechanism that connects trade and wage inequality. A large literature has argued that trade can raise inequality. Our model makes a similar prediction, but with an important twist: Trade liberalization shifts population towards larger and more productive cities, where nominal wages are higher. Thus, trade raises *nominal* wage inequality across space. To the best of our knowledge, we are the first to identify this spatial effect of trade on (nationwide) wage inequality. However, prices are also higher in larger cities. Hence, using nominal wages as the sole metric may mislead policymakers, as it exaggerates the effect of trade on inequality. Third, the model also allows us to study the reverse relationship – between urbanization and international trade. This leads to our fourth prediction: Deregulating housing supply raises nationwide exports because it leads to disproportionate growth of larger cities.

To explore the quantitative implications of our theory, we extend our simple model to a multi-location quantitative spatial model that allows for arbitrary bilateral trade costs, asymmetric coun-

³We are – to the best of our knowledge – the first to explore the implications of the thicker productivity upper tail in larger cities for the spatial distribution of exporting, and more broadly for aggregate outcomes. While the thicker upper tail in larger cities is exogenous in our model, we discuss how this feature could be endogenized via sorting (similar to Gaubert (2018)), natural advantage, or agglomeration mechanisms. The latter could be further microfounded via a search mechanism à la Chen (2024), with agglomeration leading to lower search costs in larger cities.

⁴This goes beyond what can be explained by differences in foreign market access. In fact, an important strand of the trade literature predicts the opposite pattern: A direct implication of the gravity model is that larger cities (or countries) have *lower* export intensities (Anderson and van Wincoop, 2004) because they sell disproportionately to the local (home) market.

⁵Intuitively, the thicker upper tail in larger cities implies a relatively higher mass of firms above the export cut-off. Thus, the mass of firms that enter the export market after a reduction in variable trade costs is lower relative to the mass of firms that are already exporting.

tries, heterogeneity across locations in amenities, mobility frictions due to idiosyncratic tastes for locations, as well as a richer pattern of productivity differences across space, combining elements from Redding (2016) and Melitz (2003). We structurally estimate this model using Chinese and French firm-level data. The model can account for the bulk of the correlation between export intensity and city size in the data, and it also provides a good fit for the relationship between local trade elasticities and city size. We show that for both features, the thick productivity upper tail in large cities is the crucial driving force, while differences in trade costs or in average productivity across cities play only a very limited role. To quantify the policy implications highlighted above, we perform two types of counterfactual experiments: reducing variable international trade costs and increasing the nationwide housing supply elasticity. First, moving from autarky to trade increases the spatial concentration of economic activity (measured as the share of population in the highest relative to the lowest decile of city sizes) by 11.4% (8.1%) for China (France). Second, trade benefits workers in larger relative to smaller cities as it increases real wages in the highest relative to the lowest decile of the city size distribution by 2.5% (2.7%) relative to autarky. Note that the higher welfare gains in larger cities mean that rising house prices there only partially erode the benefits of increasing nominal wages and decreasing tradable goods prices. Third, raising the housing supply elasticity nationwide from the 25% to the 75% percentile increases exports by 4.8 percent for China and 0.8 percent for France.

Lastly, we examine whether our model has quantitatively different predictions for the welfare gains from trade and increased housing supply, when compared to standard models that do not feature interactions between trade and geography. The gains from trade in our model are smaller as compared to i) a standard heterogeneous firm model without geography, and ii) a geography model with homogeneous firms. This results from the fact that exporting is concentrated in larger cities, where wages and house prices are higher, which makes it more costly for firms to expand. This dimension is missing in standard models. Hence, our unified setup has significant implications for the welfare gains from trade and urban policies.

Our paper is related to several strands of the literature. The effect of international trade on the spatial distribution of economic activity has been one of the foundational questions of economic geography. Early – mostly theoretical – contributions by Krugman and Livas Elizondo (1996), Monfort and Nicolini (2000), and Behrens, Gaigne, Ottaviano, and Thisse (2006a,b, 2007, 2009) identify different mechanisms through which trade can increase or decrease the spatial concentration of economic activity. We bring novel empirical evidence and a new mechanism to this literature. Our model marries recent advances in spatial economics (Allen and Arkolakis, 2014; Redding, 2016) with the standard trade model with heterogeneous firms (Melitz, 2003).⁶ In order

⁶The quantitative part of our paper is also related to Ducruet, Juhasz, Nagy, and Steinwender (2020), Coşar and Fajgelbaum (2016) and Fajgelbaum and Redding (2022), who use quantitative spatial equilibrium models to highlight the role of domestic transport costs for the local effects of country-level trade openness.

to replicate the observed relationship between city size and export intensity, these standard ingredients need to be combined with a thicker upper tail of the firm productivity distribution in larger cities (as documented by [Combes et al., 2012](#)). Thus, our results highlight the practical relevance of this well-known empirical pattern.

Our work also relates to a large literature on trade and inequality (see [Helpman, 2018](#), for a review). Our findings highlight that it is important to distinguish between nominal wages and welfare when examining how trade affects inequality. This results from local prices adjusting differentially across space – a feature that is missing in trade models and that was first introduced in the urban literature by [Moretti \(2013\)](#).⁷ In addition, our result that the spatial distribution of production affects the country-level trade elasticity contributes to a growing literature on the variability of this parameter ([Helpman, Melitz, and Rubinstein, 2008](#); [Novy, 2013](#); [Melitz and Redding, 2015](#); [Adão, Arkolakis, and Ganapati, 2020](#)).

Finally, we contribute to strands of the literature within international trade and economic geography. Regarding the former, we relate to the rich literature on heterogeneous firms in trade (c.f. [Bernard and Jensen, 1999](#); [Pavcnik, 2002](#); [Melitz, 2003](#)), as well as to more recent work on the heterogeneous effects of trade across locations within a country ([Autor, Dorn, and Hanson, 2013](#); [Dauth, Findeisen, and Suedekum, 2014](#); [Cheng and Potlogea, 2020](#)). In urban economics, our novel stylized fact adds to a growing literature on sorting, selection, and agglomeration across city size ([Eeckhout, Pinheiro, and Schmidheiny, 2014](#); [Davis and Dingel, 2014](#); [Gaubert, 2018](#); [Davis and Dingel, 2019](#); [Schoefer and Ziv, 2022](#)), and our policy counterfactual on housing supply constraints speaks to the literature that quantifies their importance (c.f. [Hsieh and Moretti, 2019](#)). Exploring the systematic differences in the effects of trade across cities of different sizes also relates to a growing literature on regional divergence across cities in advanced economies: [Giannone \(2022\)](#) studies the role of skill-biased technical change, [Eckert, Ganapati, and Walsh \(2022\)](#) emphasize the importance in the rise of services, ([Walsh, 2023](#)) underlines the importance of new firm creation, and [Chen, Novy, Perroni, and Wong \(2023\)](#) study the role of trade and structural transformation for the growth of French cities. We show that combining the core elements of both literatures – heterogeneity across firms and across space – leads to novel insights with important policy implications.

The rest of the paper is organized as follows: Section 2 presents the data and Section 3 our stylized fact and its robustness. In Section 4 we develop a simple model that illustrates the economic forces behind the stylized fact and explores their policy implications. Section 5 extends this to a quantitative model and performs counterfactual policy analyses. Section 6 concludes.

⁷Note that we abstract from endogenous amenity changes as highlighted by [Diamond \(2016\)](#). If those adjusted endogenously following a trade shock, there would be another margin to consider when studying the effects of trade on inequality.

2 Data

Our main empirical analysis uses firm-level data from the 2004 Chinese Economic Census of Manufacturing and from the 2000 French Unified Corporate Statistics System (FICUS). One important advantage of the Chinese and French data is that they provide detailed information on the location of firms. This allows us to study the distribution of firms and exporters across cities. In addition, we use more aggregate information at the city level from the United States (at the MSA level) and Brazil (at the microregion level) for 2012 to confirm the main patterns we derive for China and France. We begin by discussing the Chinese and French data in detail, followed by a description of the U.S. and Brazilian data. For each country, we discuss what constitutes a “city” in our data.

2.1 China

Data for the Chinese Economic Census of Manufacturing are collected by the *National Bureau of Statistics*, covering the universe of firms in China, irrespective of their size. The Chinese data contain detailed information on plant characteristics such as sales, spending on inputs and raw materials, employment, investment, and export value. In the data, the reported location of firms reflects the county where their headquarters are based. We show in Section 3.2 that this feature is unlikely to confound our results.

Our main analysis defines Chinese cities as metropolitan areas with contiguous lights in night-time satellite images. We use the correspondence constructed by [Dingel, Miscio, and Davis \(2019\)](#) to map counties into metropolitan areas with a threshold for light intensity equal to 30.⁸ This value is in the middle of the set of thresholds provided by these authors. Importantly, our results do not depend on the particular light intensity threshold. For each metropolitan area, we use information on the urban population of the underlying counties, which is provided by the Chinese Population Census of 2010 (i.e., the Chinese Census distinguishes between rural and urban population within each county). We then define ‘city size’ as the aggregate urban population of the Metropolitan Area.⁹

The Chinese Census of Manufacturing contains information for approximately 1,240,000 firms located in cities to which we can match population information in 2004. We drop firms with zero or missing sales (67,286 observations, corresponding to 5.4% of the sample), non-manufacturing or missing industry codes (124,310, 10.0% of the sample), or export intensity above 100% (5,397

⁸Counties are the third administrative division in China, below provinces and prefectures – the other administrative division researchers typically use to define cities (e.g., [Brandt, Van Biesebroeck, and Zhang, 2014](#)). For reference, large metropolitan areas (urban population over 1 million) have, on average, about nine counties. In contrast, most small metropolitan areas (population below 100,000) comprise a single county.

⁹A large body of research using information from China defines cities in terms of prefecture-level cities. A prefecture-level city is an integrated political and economic unit, often including rural areas. We avoid defining cities in terms of prefectures because administrative boundaries may fragment economically integrated areas into distinct cities or circumscribe places, including rural areas.

observations, 0.4% of the sample). We also drop processing trade (9,672 observations, defined as firms where processing exports account for over 90% of their sales). In addition, to ensure meaningful variation in export intensity at the city level, we only consider cities with at least 250 firms. Our final sample consists of 916,870 firms in 629 cities (metropolitan areas).

2.2 France

Our analysis for France uses firms from the 2000 Unified Corporate Statistics System (FICUS). FICUS is an administrative data set collected by the French National Statistical Institute (*Institut National de la Statistique et des Études Économiques*, INSEE), covering the universe of private sector firms. It reports information on domestic and export revenue, industry classification, headquarters location (commuting zones), employment, capital, value-added, and production.

We define cities in France in terms of commuting zones (*zone d'emploi*). We use the definition of commuting zones published by INSEE in 2011 that assigns municipalities (“communes”) to commuting zones (“zone d'emploi”) to create “*geographical area[s] within which most of the labour force lives and works.*”¹⁰ City size reflects the overall commuting zone population, which we obtain by aggregating municipality-level information from the French Population Census of 1999. As in the case of China, we restrict the analysis to firms with positive information on exports and sales and to cities with at least 250 firms. The final sample contains all 304 commuting zones in mainland France.

2.3 United States

In the case of the United States, we define cities in terms of Metropolitan Statistical Areas (MSA). MSAs are defined by the United States Office of Management and Budget as one or more adjacent counties with at least one urban area with a population of at least 50,000 inhabitants and characterized by a high degree of social and economic integration, as measured by commuting flows to work and school. As Dingel et al. (2019) show, MSAs are well-approximated by cities defined in terms of contiguous areas of lights in nighttime satellite images, making the city definition comparable to China. Our analysis considers 324 U.S. metropolitan areas with a population over 100,000 inhabitants in 2012.

To develop our main analysis, we combine data from several sources.¹¹ Data for exports at the MSA level are provided by the International Trade Administration of the U.S. Department of Commerce and include overall exports.¹² We combine this with establishment-level information of sales and employment aggregated at the MSA level from the 2012 Economic Census.¹³ Con-

¹⁰<https://www.insee.fr/en/metadonnees/definition/c1361>.

¹¹Most U.S agencies provide tabulations on key economic accounts at the MSA level. This contrasts with China, where we aggregate county-level information to derive statistics for metropolitan areas.

¹²<https://www.export.gov/Metropolitan-Trade-Data>.

¹³<https://www.census.gov/programs-surveys/economic-census/data/datasets.2012.html>

sequently, city-level export intensity is constructed as overall exports over sales across all sectors. Finally, we use MSA population from the population projections of the U.S. Census Bureau.¹⁴

2.4 Brazil

Finally, in the case of Brazil, we consider microregions as the main unit of analysis. Microregions are defined by the Brazilian Institute of Geography and Statistics (IBGE) as urban agglomerations of economically integrated contiguous municipalities with similar geographic and productive characteristics.¹⁵ Although microregions do not directly capture commuting flows (in contrast to U.S. Metropolitan Areas, or French employment zones), they are constructed according to information on integration of local economies, which is closely related to the notion of local labor markets. Our sample includes 420 microregions with more than 100,000 inhabitants in 2012.

We use municipality-level overall exports from the COMEX Stat database (compiled by the Brazilian *Ministry of Industry, Foreign Trade and Services*).¹⁶ We complement this data source with municipal-level GDP from IBGE.¹⁷ We aggregate both exports and GDP from the municipality level to the level of microregions using the correspondence provided by the IBGE, and compute city-level export intensity as the ratio of overall exports over GDP. Finally, we use population projections from the 2010 population Census to measure city (microregion) size.¹⁸

2.5 Summary Statistics and Export Intensity

Before turning to our empirical results, we show descriptive statistics for the sample of cities considered in the analysis for China, France, the United States, and Brazil. Table 1 shows statistics for the distribution of population and export intensity for the four samples. Average city size varies importantly across the datasets. U.S. cities are larger on average (about 800,000 inhabitants), followed by China (710,000), Brazil (464,000), and France (193,000). These reflect the fact that the population in the U.S. is more concentrated in larger cities.¹⁹ While for the U.S., two-thirds of the cities in our sample have populations above 500,000, in China and Brazil, 16 percent of the cities surpass this threshold, and in France, only 6 percent.

We define export intensity as the share of a city's sales that are exported. Correspondingly, we

¹⁴<https://www.census.gov/data/tables/2012/demo/popproj/2012-summary-tables.html>

¹⁵A number of researchers have used microregions as their main unit of analysis (see Kovak, 2013; Dix-Carneiro and Kovak, 2015, 2017b, 2019; Costa, Garred, and Pessoa, 2016; Chauvin, Glaeser, Ma, and Tobio, 2017).

¹⁶The COMEX Stat database can be publicly accessed through an interactive interface in the website <http://comexstat.mdic.gov.br>. In our analysis, we downloaded the version compiled by the Ministry of Economics from the information in the COMEX Stat database at the municipal level (<https://www.gov.br/produktividade-e-comercio-exterior/pt-br/assuntos/comercio-exterior/estatisticas/base-de-dados-bruta>).

¹⁷<https://www.ibge.gov.br/en/statistics/economic/national-accounts/19567-gross-domestic-product-of-municipalities.html>

¹⁸<https://www.ibge.gov.br/en/statistics/social/education/18391-2010-population-census.html?=&t=microdados>.

¹⁹This is consistent with evidence in Au and Henderson (2006), who show that about half of prefecture-level cities in China are smaller than their optimal size. They argue that this is most likely due to the existence of strong migration restrictions.

define the export intensity in city i (ρ_i) as follows:

$$\rho_i = \frac{x_i}{r_i}, \quad (1)$$

where x_i and r_i denote city-level exports and revenues, respectively (i.e., aggregated across all firms operating in city i). The distribution of export intensity is positively skewed for all countries in our sample, with a substantially fatter upper tail in Brazil and China than in France and the United States (see Figure A.1). In the U.S. and France, all cities in our sample have exporting firms; in contrast, about 2 and 6 percent of the cities record no export activity in China and Brazil, respectively. While noteworthy, the presence of cities with zero exports does not affect the magnitude of our results because these cities represent a small fraction of output (0.3% and 0.9% of the production in China and Brazil, respectively).

3 Empirical Results: Export Activity and City Size

This section presents our main empirical results. Using data from China, France, the United States, and Brazil we find that export activity is concentrated in larger cities. We show that this pattern (i) is predominantly driven by within-industry variation, (ii) holds when using different definitions of export activity and city size, and (iii) is not driven by manufacturing alone. We then provide evidence that heterogeneity in firm productivity is an important underlying mechanism.

3.1 Baseline Results: The Stylized Fact in Four Countries

We run the following regression to establish our core stylized fact:

$$\ln(\rho_i) = \alpha + \beta \ln(pop_i) + \gamma X_i + \varepsilon_i, \quad (2)$$

where ρ_i is the export intensity of city i , as defined in (1), pop_i is city population, and X_i is a vector of several proxies for domestic and international trade costs: Average distance to other domestic cities, distance to the border, distance to the coast, border dummies, and coastal dummies.²⁰ Our coefficient of interest is β , reflecting the elasticity between export intensity and city size. Table 2 presents the results: We obtain statistically highly significant estimates for all four countries, ranging from 0.16 for the U.S. to 0.34 for China. Thus, doubling city size in China is associated with raising export intensity by about one-third. With an average export intensity of 8.8%, this corresponds to an increase in the fraction of exported (relative to total) local output by about 3 percentage points. Importantly, the coefficients remain quantitatively similar and statistically highly significant when we include the geographical controls X_i (even columns).

Since our empirical findings are robust to controlling for (domestic and foreign) market access,

²⁰For a city, a border dummy takes on value 1 if the city boundaries overlap with a border segment.

they cannot simply be explained by larger cities emerging in locations with lower international trade costs. Figure 1 provides binscatter plots for the regressions from even columns in Table 2, illustrating that our results are not driven by outliers.

3.2 Robustness Checks

We implement several tests to check the robustness of our main finding for China and France, where we have the most detailed data. We summarize these results here; the corresponding tables and figures are shown in Appendix A.

Alternative measure of export intensity. Our baseline export intensity measure uses city-level exports relative to sales. While this is the natural normalizing variable, we present results with an alternative denominator – city population. Using this variable, we repeat our main analysis with *per capita* exports as a proxy for export activity. Figure A.2 plots per-capita exports against city size, and Table A.1 shows the corresponding elasticities. All coefficients are statistically significant at the 1 percent level, with magnitudes similar to our baseline results.

Population density. Our baseline specification uses city population as the main explanatory variable. An alternative measure of agglomeration that is widely used in the literature is population density (see for instance Combes et al., 2012). This measure can deviate substantially from city size, especially when cities vary widely in the size of their geographic areas.²¹ Panel A of Table A.2 shows that our results hold when we use population *density* as the explanatory variable, and Figure A.3 visualizes the relationship between export intensity and population density (both in logs). Finally, we also obtain similar results when using per-capita exports as dependent variable in combination with population density (see Panel B of Table A.1 and Figure A.4).

Export intensity in manufacturing vs. services. Our baseline analysis for China only considers production and export information for the manufacturing industry. The French data allow us to estimate (2) separately for manufacturing, services and the primary sector. Table A.3 reports the results, show that the gradient of export intensity with city size is remarkably similar for firms in manufacturing and in services. We also provide a placebo check, showing that export activity of the primary sector – which depends on land abundance and natural resources – is indeed *less* concentrated in larger cities.

China-specific robustness checks. We perform a number of robustness checks for our results for China in the appendix. Table A.4 shows that our results are not affected by i) controlling for Special Economic Zones (SEZ) and Coastal Development Areas (CDA), ii) by using only direct exports when computing export intensity, iii) when using prefecture-level Chinese cities as the main unit of analysis (e.g., Au and Henderson, 2006).

Multi-location firms. For our firm-level data (France and China), firm location is only defined at

²¹See Henderson, Nigmatulina, and Kriticos (2021) for a recent study discussing the power of different measures to estimate urban agglomeration effects in the context of six Sub-Saharan African countries.

the headquarter-level. This may introduce an upward bias if export-intensive companies with production based in small cities locate their headquarters in large cities.²² As we describe in Appendix A.3, we can directly address this concern for France, using administrative data on employment and wages by establishment. We show in Table A.5 that i) elasticities are very similar when restricting the analysis to the sample of firms that are active in a single commuting zone (Panel A), and ii) for firms with establishments in multiple commuting zones, our results hold when we assign domestic and export revenue across each firm’s establishments proportionally to the respective wage bill (Panel B). For China, we build on the fact that fewer than 10 percent of firms are multi-plant firms in the Chinese manufacturing industry, and these tend to be relatively large (Brandt et al., 2014). Thus, excluding large firms is a way to (indirectly) control for the role of multi-plant firms. Table A.6 shows that the estimated elasticity is very similar when dropping the top 1%, 5%, or 10% of the overall or within-sector employment distribution. We conclude that our baseline results are unlikely to be driven by export-intensive companies locating their headquarters in large cities.

3.3 2SLS Results

An important concern with our empirical results is that unobserved endogenous factors may drive the relationship between export intensity and city size. For example, larger cities may have invested disproportionately in export-enabling infrastructure, or there may even be a reverse channel, with high exporting activity raising firm efficiency (c.f. Garcia-Marin and Voigtländer, 2019) and leading to larger city size. To further address this issue, we implement a two-stage least squares (2SLS) strategy for our two main countries of analysis – China and France. We use historical city population from the 16th and 19th century, respectively, as an instrument for modern population. The underlying assumption is that historical population was not determined by the same factors that affect city size and export intensity today. This assumption is reasonable, given the significant changes in transport technology, industry composition, and policy over the last centuries. Trade patterns have changed significantly over this period, weakening the link between current and historical exporting. As in our baseline specification, we control for a number of time-invariant geographic features associated with trade costs, such as location on the coast and distance to the border.

For China, we instrument current population using prefecture-level population in the 1580s from Bai and Jia (2021).²³ For France, we use population records for 1876 as an instrument for the

²²This problem is not present in the U.S. data, where exports are aggregated to the city level from the establishment level. For Brazil, the data is aggregated from the tax domicile of the company, which can be at the establishment or the firm level, depending on how the firm files taxes.

²³Firms in our data are located in 344 prefectures. However, Bai and Jia (2021) only provides historical population information for the 260 prefectures belonging to China proper – i.e., excluding the remaining 74 prefectures from Inner China provinces, such as the Tibet or Inner Mongolia provinces. Our 2SLS regressions for China use the 260 prefectures with available data. Table A.4 provides OLS results for the full set of prefectures. See Bai and Jia (2021) for details on the mapping of historical population records to current prefecture geographical areas. We recompute our

current population. INSEE recently published this information at the level of current municipalities, thus avoiding issues related to changing municipality borders.²⁴ We aggregate the municipal population records to the commuting zone level, using the official correspondences provided by INSEE.²⁵

Table 3 presents our 2SLS results on export intensity and city size. We begin by describing the results for China (columns 1-4). Column 1 confirms that the elasticity from the baseline OLS regression remains similar in magnitude and highly significant when using prefectures instead of metropolitan areas to define cities. Column 2 presents results from a reduced-form specification, where we directly regress contemporaneous export intensity on the 1580s prefecture-level population. We find a strong positive relationship between the two variables. Next, in column 3, we report the first-stage results, where we regress contemporaneous urban population on historical prefecture-level population. We obtain a strong first stage, with an F-statistic substantially above the Stock-Yogo critical value of 16.4 for 10% maximal IV bias. We also note that the coastal prefecture dummy is the only statistically significant control; no other geographical variable is significant in the first-stage regression. The estimated first stage coefficient implies that a 10% higher historical population in 1580 predicts a 3% larger population today. Finally, column 4 shows the second-stage result. The estimated coefficient on city size is statistically significant and remarkably similar to the OLS coefficient in column 1.

Columns 5-8 in Table 3 present our 2SLS results for France. The reduced-form coefficient is similar to the one for China. The first-stage coefficient is stronger, which is likely due to the shorter time gap between the instrument and contemporaneous city size. Finally, the 2SLS coefficient (column 6) is statistically significant and similar to its OLS counterpart.

An important caveat with our 2SLS results is that the exclusion restriction may be violated if historical city size is associated with unobservables that also affect contemporaneous exporting. Our controls for market access can only partially alleviate this concern, so that our 2SLS results need to be interpreted with caution. We are, however, confident that these results can address reverse causality (exporting fostering firm productivity and city size) as well as endogenous investments in modern infrastructure (which could drive exporting and may vary systematically with city size).

3.4 Mechanisms

In what follows, we shed light on the mechanism that drives our novel stylized fact. We build on prominent features of existing trade models to motivate two decompositions that distinguish our new-new-trade-theory mechanism from alternative explanations.

control variables using the location and borders of prefecture-level cities, and link them to the prefecture-level dataset.

²⁴*Historique des populations communales: Recensements de la population 1876-2019*, available at <https://www.insee.fr/fr/statistiques/3698339>.

²⁵<https://www.insee.fr/fr/information/2114596>

Within- and between-industries variation. Theories of comparative advantage could explain our stylized fact through differences across industries, while new-new-trade-theory-based mechanisms would generate variation in export intensity *within* industries, across firms. To distinguish between these mechanisms, we decompose export intensity into its underlying within- and between-industry components (see Appendix A.4 for detail). The between-industry component is the counterfactual measure that would result if city-level export intensity only varied due to differences in industry composition across cities.²⁶ The within-industry component is the residual variation that is not accounted for by differences in the sectoral composition of cities.

Table 4 shows the results of the decomposition for our main datasets, China and France. Note that by construction, the within- and between-industry coefficients add up to the overall elasticity between export intensity and city size from Table 2. Columns 3 and 6 in Table 4 report the share of the overall variation that is accounted for by the within-industry component – 90% and 52% for China and France, respectively. Thus, for China, the relationship between export intensity and city size is almost exclusively driven by differences within industries. For France, differences across industries play a larger role. This is likely due to the fact that the French data comprise all sectors, including many with minimal exporting such as construction (recall that the Chinese data only include manufacturing). In fact, if we restrict the sample to manufacturing in France, the within-component accounts for 67%. Thus, within-industry differences account for the majority of our stylized fact in both countries.²⁷ In sum, for both countries, the majority of the variation is driven by systematic differences in exporting behavior across firms *within* the same industry.

Intensive margin and differences in transport costs. Next, we check whether a standard gravity model could explain our stylized fact through differences in variable trade costs. In particular, could our stylized fact be driven by larger cities facing systematically lower variable trade costs (for example, because they are more likely to be located on the coast or benefit from larger endogenous infrastructure investments)? In order to test for this possible mechanism, we decompose export intensity into the intensive margin (export revenue over the total revenue of exporters: x_i/r_i^x) and the remainder (the total revenue of exporters over the total revenue of all firms r_i^x/r_i):

$$\ln \left(\frac{x_i}{r_i} \right) = \ln \left(\frac{x_i}{r_i^x} \right) + \ln \left(\frac{r_i^x}{r_i} \right) \quad (3)$$

In gravity models, differences in variable transport costs would affect export intensity through the

²⁶We use detailed, 4-digit industries in our decomposition. The decomposition cannot be performed for Brazil and the United States, as for these countries we only have access to aggregate city-level exports (i.e., not by sector).

²⁷Another way to gauge the importance of the industry dimension is to run our main regression (2) at the *city-industry* level with and without industry fixed effects. Table A.7 in Appendix A.4 shows that adding industry fixed effects does not change our core result for China, while it reduces the city-size – export intensity coefficient by about one-third for France. We prefer our main specification at the city level as it avoids issues with the presence of zeros at the city-industry level and allows for the entire analysis to be conducted at the same level of aggregation.

intensive margin.²⁸ We thus focus on this component in our analysis but also report results for the remainder.²⁹ In Table 5, we regress these components on city size and geographical controls. Since this is an exact decomposition, the coefficients add up to our baseline result (reported again in columns 1 and 4). The intensive margin is negatively correlated with city size in China (column 2) and uncorrelated with city size in France (column 5).³⁰ For both countries, we find an economically large and statistically significant correlation between city size and the importance of exporters in local economic activity (r_i^x/r_i). Most importantly, however, the negative or insignificant role of the intensive margin implies that our stylized fact is not driven by systematically lower variable trade costs in larger cities.

In sum, the results of the two decompositions suggest that our stylized fact is driven by differences across firms *within* industries that are unrelated to differences in variable trade costs. This leaves systematic differences in productivity across small and large cities (as first documented by Combes et al., 2012) as the natural mechanism behind our stylized fact.

4 Simple Model

In this section, we present a simple model to highlight the key mechanism that can explain our novel stylized fact. The model also serves as the foundation for our quantitative analysis in Section 5. It incorporates firm heterogeneity (Melitz, 2003) driven by differences in the upper tail of the local productivity distribution into an open economy Rosen-Roback framework. In the model, locations only differ in their exogenous firm productivity distribution, which we assume to be Pareto with location-specific shape parameters. Domestic and export market entry of firms, together with labor mobility across locations, give rise to the equilibrium distribution of economic activity, where productive locations i) host larger cities and ii) are more export-intensive.

4.1 Setup

We consider a world economy featuring 2 symmetric countries.³¹ Each country is endowed with an exogenous population L of identical workers and an exogenous number of locations $i = 1, \dots, I$.

²⁸It is important to note that for competing models without firm heterogeneity to match our main stylized facts, it is not sufficient for larger cities to have systematically better foreign market access. What is required instead is that larger cities have systematically better foreign *relative to domestic* market access than smaller cities. This, in turn, would lead to a higher intensive margin of exporting in larger cities.

²⁹The remainder can be interpreted as the importance of exporters in economic activity, reflecting both differences in export participation and in the size (revenues) of exporting firms across cities. We do not disentangle the term into these sub-components, as it would not lead to additional insights on mechanisms.

³⁰The negative coefficient for China can be rationalized as follows: Larger cities have larger domestic market access due to the larger local home market. This will lead c.p. to *lower* exports relative to domestic sales (i.e., lower export intensity). Thus, to account for our stylized fact, any differences in international variable trade cost would have to outweigh and overcompensate the home market effect for larger cities. Our results suggest that this is far from being the case in China, while some of the negative home market effect is absorbed in France.

³¹We therefore suppress country subscripts in the description of the model.

Within countries, all locations are endowed with the same stock of local land N . In each location i , a city with endogenous population L_i emerges. Workers are assumed to be perfectly mobile across locations within countries but immobile internationally. In order to focus on the role of differences in the productivity distribution across locations, our simple model abstracts from within-country trade costs and differences in amenities.

Preferences. Workers are homogeneous; they decide where to locate and consume a bundle of goods and housing, while earning the local wage w_i in city i . Worker preferences are given by:

$$U_i = c_i^\beta h_i^{1-\beta} \quad (4)$$

where h_i denotes (individual) housing consumption in city i , and c_i is a CES composite of the consumption of tradable varieties:

$$c_i = \left[\int c_i(x)^{\frac{\sigma-1}{\sigma}} dx \right]^{\frac{\sigma}{\sigma-1}} \quad (5)$$

In each city, housing is produced by atomistic housing developers j , using land and capital according to the technology:

$$h_i(j) = k_i(j)^{1-\gamma} n_i(j)^\gamma \quad (6)$$

where $k_i(j)$ and $n_i(j)$ denote capital and land used by developer j at location i . Both land and housing markets are assumed to be perfectly competitive at the local level. Capital is provided in perfectly elastic supply at the exogenous rental rate $r_K = 1$, while each location i has the same land endowment N . Land market clearing implies $\int_j n_i(j) = N$. Total housing supply at the city level is then given by

$$H_i = K_i^{1-\gamma} N^\gamma \quad (7)$$

where $K_i = \int_j k_i(j)$. Land and capital income are assumed to be fully taxed by local governments and redistributed lump-sum to local residents as in [Helpman et al. \(1995\)](#).

Production. In each country and city, there is an unlimited supply of potential entrants. Each firm produces a differentiated tradable variety using local labor. Once they've chosen a city to enter and paid the entry cost F_e in units of local labor, firms draw their productivity ψ from the location-specific productivity distribution $g_i(\psi)$. In each location, productivity is distributed Pareto with a common scale parameter set to value 1 and a location-specific shape parameter α_i . The production technology is given by:

$$q = \psi l \quad (8)$$

where l denotes the local labor input required to produce a quantity of output q for a firm of productivity ψ . Depending on their productivity, firms choose to exit immediately, serve the domestic

market paying fixed cost F_d (in units of local labor), or serve both the domestic and the export market (paying an additional fixed cost F_x to export).³² All fixed costs as well as the variable costs of exporting (τ) are identical across locations. We adopt the standard simplifying assumption of zero domestic transport costs.

Firms engage in monopolistic competition. Given CES demand, this implies that prices are set at a constant mark-up over marginal cost. Because firms can serve all domestic locations without incurring variable trade costs, and because imported varieties also reach each location at the same international trade cost, the price index of the tradable good is the same across locations. We choose $P = 1$ as the numeraire.

For a firm with productivity draw ψ , operational profits from serving the domestic and foreign market, respectively, are given by:

$$\pi_i^d(\psi) = \frac{r_i^d(\psi)}{\sigma} - w_i F_d = \frac{R}{\sigma} \left(\frac{\psi \rho}{w_i} \right)^{\sigma-1} - w_i F_d \quad (9)$$

$$\pi_i^x(\psi) = \frac{r_i^x(\psi)}{\sigma} - w_i F_x = \tau^{1-\sigma} \frac{R}{\sigma} \left(\frac{\psi \rho}{w_i} \right)^{\sigma-1} - w_i F_x \quad (10)$$

where $r_i^d(\psi)$ and $r_i^x(\psi)$ denote the revenue from serving the domestic and foreign market. Finally, we define $\rho \equiv \frac{\sigma-1}{\sigma}$, R denotes the aggregate revenue of the tradable sector in each country, and w_i is the wage in location i .

As in Melitz (2003), the industry equilibrium (in our model, for each location) can be characterized using two zero-profit cutoff conditions (11) that yield expressions for the minimum productivity threshold for successful entry into the domestic market (ψ_i^{d*}) and for successful entry into the foreign market (ψ_i^{x*}), as well as a free entry condition (12). The latter states that the expected operational profits of market entry $\bar{\pi}_i$ are equal to the fixed cost of market entry (in terms of local labor).

$$\pi_i^d(\psi_i^{d*}) = 0 \quad \text{and} \quad \pi_i^x(\psi_i^{x*}) = 0 \quad (11)$$

$$\bar{\pi}_i = \int_{\psi_i^{d*}}^{\infty} \pi_i^d(\psi) g_i(\psi) d\psi + \int_{\psi_i^{x*}}^{\infty} \pi_i^x(\psi) g_i(\psi) d\psi = w_i F_e \quad (12)$$

We use (12) to solve for the measure of entrants (M_i^e) in each city:³³

$$M_i^e = \frac{L_i}{\sigma \left[F_e + \psi_i^{d*- \alpha_i} F_d + \psi_i^{x*- \alpha_i} F_x \right]} \quad (13)$$

³²We impose the parametric restriction that $\tau^{\sigma-1} F_x > F_d$, which ensures that firms serving the foreign market represent a strict subset of the firms that operate in the domestic market.

³³ M_i^e represents the measure of firms that pay the fixed cost of entry F_e at each location. This measure includes both active firms (that serve at least one market) and inactive firms (i.e., firms with 0 total revenues) in each city.

The thick upper tail assumption. A key ingredient of our model is the thicker upper tail of the productivity distribution $g_i(\psi)$ in larger cities. This assumption reflects an empirical pattern in our data that has previously been documented by [Combes et al. \(2012\)](#). We show below in Section 5.2 that the same pattern holds in our data. Our model takes the differential fat-tails as given and explores their *consequences* for the spatial distribution of exporting and other macroeconomic outcomes. Before continuing with the model setup, we note that these city-level productivity distributions could be readily endogenized, albeit at some cost in terms of tractability and notational complexity.

For instance, we could assume that intrinsic firm productivity distributions are identical across space but that the realized firm-level productivity also depends on location-specific characteristics (“natural advantages”): $\ln \psi_i(z) = \theta_i \ln(z)$, where θ_i is a location- i -specific productivity shifter and z are firm-level productivity draws from a standard Pareto distribution with shape parameter α that is identical across locations. Such a formulation would be equivalent to our model, with the resulting city-level productivity distributions being standard Pareto with shape parameters $\alpha_i = \alpha/\theta_i$. We could further endogenize θ_i by introducing agglomeration economies: $\theta_i = \theta(L_i)$, where the productivity shifter $\theta(\cdot)$ is increasing in city population L_i . Such a formulation would capture a tight connection between productivity, city size, and the shape of city-level productivity distributions. However, it would likely come at the cost of multiple equilibria.

Both the natural-advantage and agglomeration-economies-based accounts can be further micro-founded via search mechanisms such as the one developed in [Chen \(2024\)](#). Here, exogenous or agglomeration-driven differences in search costs across space could ground differences in the thickness of the upper tail: Locations with lower search costs would feature firm productivity distributions with thicker upper tails.³⁴

Residential choice in spatial equilibrium. In spatial equilibrium, utility is equal across all locations within each country and given by:

$$U_i = \beta^\beta (1 - \beta)^{1-\beta} \frac{v_i}{(p_i^h)^{1-\beta}} = \bar{U} \quad \forall i \quad (14)$$

where v_i denotes total income per worker, and p_i^h is the price of housing at location i . Since we assume that capital and land income are fully taxed by local governments and redistributed lump

³⁴Alternative accounts that yield thicker upper tails of firm productivity distributions in more productive locations (albeit typically not of the tractable Pareto form) involve sorting mechanisms (similar to those in [Gaubert \(2018\)](#) or [Behrens, Duranton, and Robert-Nicoud \(2014\)](#)) where at least some firms (or entrepreneurs) are mobile across cities within countries and free to sort into the most desirable locations.

sum to workers, total income in location i is given by labor income plus expenditure on housing:

$$v_i L_i = w_i L_i + (1 - \beta) v_i L_i = \frac{w_i L_i}{\beta} \quad (15)$$

The price of housing follows from (7) and is given by:

$$p_i^h = \left(\frac{r_K}{1 - \gamma} \right)^{1-\gamma} \left(\frac{r_i}{\gamma} \right)^\gamma = \left(\frac{1}{1 - \gamma} \right)^{1-\gamma} \left(\frac{r_i}{\gamma} \right)^\gamma \quad (16)$$

where r_i denotes the (endogenous) land rental rate at location i , and $r_K = 1$ is the (given) rental rate of capital.

Finally, land market clearing yields the equilibrium land rent:

$$r_i = \frac{1 - \beta}{\beta} \gamma \frac{w_i L_i}{N} \quad (17)$$

General equilibrium. The general equilibrium of the model is represented by a measure of workers L_i , a set of entry thresholds for the domestic and the export markets, ψ_i^{d*} and ψ_i^{x*} , respectively, a set of wages w_i for each location, as well as the level of utility in the economy \bar{U} that satisfy the following system of equations:

1. The industry equilibrium consisting of the zero profit cut-off conditions and the free entry condition (equations 11 and 12)
2. Trade balance at each location (wage income in each location equals expenditure on goods produced in that location)

$$w_i L_i = M_i^e \left(\int_{\psi_i^{d*}} r_i^d(\psi) g_i(\psi) d\psi + \int_{\psi_i^{x*}} r_i^x(\psi) g_i(\psi) d\psi \right) \quad \forall i \quad (18)$$

3. Spatial equilibrium for workers among all domestic locations

$$U_i = \bar{U} \quad \forall i \quad (19)$$

4. National labour markets clear

$$\sum_i L_i = L \quad (20)$$

4.2 Matching the Stylized Fact

We now show that our simple model can match the novel stylized fact that we documented in the data. We begin with an auxiliary result that will facilitate our exposition.

Lemma 1. *Cities whose exogenous productivity distributions feature a thicker upper tail (i.e., smaller shape parameters α_i) will have higher average firm productivity in equilibrium.*

Proof: See Appendix C.1.

The exogenous productivity distributions in cities with a smaller shape parameter feature a fatter upper tail and thus more firms with high productivity draws. This ex-ante difference is reinforced by the endogenous selection into entry, since cities with more productive firms feature stronger labor demand for a given wage and hence stronger selection effects. We will therefore refer to cities with smaller shape parameters simply as “more productive” cities. We now proceed to stating our first theoretical result, connecting city size to export intensity

Proposition 1. *In equilibrium, more productive cities:*

- (i) *Have higher population and higher export intensity.*
- (ii) *The higher export intensity is driven by the proportion of exporters among local firms (i.e., by the extensive as opposed to the intensive margin of exporting).*

Proof: See Appendix C.2.

Intuitively, the presence of highly productive firms implies higher local labor demand, so these locations feature higher equilibrium wages, population, and house prices. Crucially, more productive cities also have a higher export intensity. This is in contrast to standard models that only feature differences in the *level* of productivity across locations (e.g., Redding, 2016): As in Melitz (2003), the productivity thresholds for serving the domestic and the export market are proportional to each other in our model, and the ratio is common to all locations:

$$\frac{\psi_i^{x*}}{\psi_i^{d*}} = \tau \left(\frac{F_x}{F} \right)^{\frac{1}{\sigma-1}} \quad \forall i \quad (21)$$

Thus, standard models that build on differences in *average* productivity across locations do not feature spatial heterogeneity in export intensity. In our setting, the fatter upper tail of the productivity distribution in more productive cities means that for a given ratio in (21), a higher share of firms are above the export threshold. This, in turn, leads to more productive locations having a higher export intensity. Note that this is driven exclusively by the export share (i.e., the fraction of firms that export), with the intensive margin of exporting (i.e., the export intensity of exporters) being constant across space.³⁵

Taken together, our simple model can account for our novel stylized fact and for the result from the decomposition in Section 3.4, showing that it is driven by the extensive margin of exporting.

³⁵This feature of the model is a direct result of the fact that under CES preferences (as in Melitz, 2003), conditional on exporting, firm-level export intensity is constant with respect to firm productivity.

The crucial feature in our model that delivers these predictions is the thicker productivity upper tail in larger (more productive) cities. This pattern has been empirically documented by Combes et al. (2012) for France, and it also holds in our data; but its implications have thus far not been exploited theoretically.

4.3 Comparative Statics

Our model provides novel insights into the joint determination of international trade and economic geography outcomes. In this section, we highlight these interactions by studying the impact of trade and spatial policies on a variety of variables.

Implications of international trade liberalization for internal geography. In our simplified setting, the spatial reallocation of employment associated with trade liberalization is straightforward:

Proposition 2. *A reduction in the variable international trade costs τ leads to a shift in population towards larger cities, as well as an increase in aggregate productivity and welfare.*³⁶

Proof: See Appendix C.3.

Trade liberalization leads to two types of reallocation of economic activity: First, within each location, economic activity reallocates from less to more productive firms. This raises the location-specific average productivity and puts upward pressure on wages. Second, this effect is stronger in larger cities due to their higher initial export intensity. Thus, there is a reallocation of workers to larger cities. Both of these channels increase aggregate productivity and welfare.

Trade costs and inequality. Since we focus on equilibrium outcomes with mobile, homogeneous workers, there are no distributional effects of trade in terms of utility. However, many studies on the distributional effects of trade examine wage inequality. We thus highlight how our mechanism naturally gives rise to both regional nominal wage inequality (i.e., wage inequality across cities) and aggregate wage inequality among workers (e.g., the nation-wide Gini coefficient). These results highlight the importance of accounting for geography and spatial equilibrium when studying the effects of trade on inequality. Formally:

Proposition 3. *A reduction in the variable international trade costs τ leads to*

- (i) *an increase in nominal wage inequality across cities, with nominal wages in larger cities increasing relative to smaller cities. However, trade liberalization does not lead to welfare inequality across cities.*
- (ii) *a change in nominal wage inequality at the country level. The direction of this effect is ambiguous and depends on the interplay of the wage changes from (i) and population re-*

³⁶This proposition refers to the long run, with perfectly mobile labor. If labor is immobile, a reduction in τ will instead lead to differences in welfare across space, with workers in small cities achieving a lower level of utility than workers in large cities.

allocations. However, trade liberalization does not lead to a change in aggregate welfare inequality.

Proof: See Appendix C.4.

We already established that trade liberalization disproportionately increases the demand for labor in larger and more productive cities. This not only leads to disproportionate population growth in more productive cities (Proposition 2) but also pushes up their wages, as they need to attract workers while compensating for rising house prices. Additionally, less productive cities lose population, and housing becomes cheaper. Thus, wages in these places decline relative to larger cities. Consequently, trade liberalization increases the pre-existing nominal wage gap between small and large cities. However, due to perfect worker mobility, this rise in nominal wage inequality does not imply inequality in *utility* in spatial equilibrium. This is similar in spirit to Moretti (2013) and Diamond (2016), highlighting the importance of getting the inequality measure right.

Aggregate nominal wage inequality in the model depends on differences in nominal wages across cities and on the distribution of population across cities. Trade liberalization increases nominal wage differences across cities as outlined above (Proposition 3, i) and also changes the distribution of population across space (Proposition 2). Since the effect of the latter on aggregate wage inequality is ambiguous, we cannot sign this effect and will explore it more thoroughly below in the quantitative model.

Spatial policy and trade. The proposition below characterizes the implications of an increase in housing supply elasticities for international trade:

Proposition 4. *Planning policies that increase the housing supply elasticity lead to an increase in country-level export intensity and aggregate productivity.*

Proof: See Appendix C.5.

Intuitively, relaxing housing supply restrictions reduces the cost of housing in all locations, thus making all cities more attractive. However, this effect is stronger for larger and more productive cities, as these locations were more constrained by the scarcity of land and housing. As a result of this mechanism, workers move towards the most productive cities. Thus, the relative mass of firms shifts to the more productive locations, where firms are also more likely to be exporters. This process causes an increase in aggregate productivity and in aggregate export intensity.

Local and aggregate trade elasticity. Our simple model makes two novel predictions regarding a key object of interest in the international trade literature: the trade elasticity. First, the model predicts that city-level international trade elasticities (i.e., the elasticity of city-level exports with respect to the variable foreign trade cost τ) are heterogeneous across cities and systematically lower in larger cities. Second, the aggregate (or country-level) trade elasticity depends on the

endogenous distribution of exporters across space. To derive these predictions formally, we make the simplifying assumption that each city is small relative to the aggregate economy:

Proposition 5. *Assume each city is small relative to the national economy and that for each city, foreign demand is small relative to domestic demand (i.e., τ and/or F_x are large). Then:*

- (i) *each city's (international) trade elasticity is given by the shape parameter of its firm productivity distribution.*
- (ii) *the city-level (international) trade elasticity is decreasing with city size.*
- (iii) *the country-level trade elasticity is endogenous and given by an export-weighted average of city-level trade elasticities.*

Proof: See Appendix C.6.

Under the small open economy conditions at the city level and the condition of relatively small effective foreign demand, the response of city-level exports to a decline in international trade costs is similar to that in Chaney (2008). The response of exports to changes in trade costs can be decomposed into an intensive-margin (equal to $\sigma - 1$) and an extensive-margin-response (equal to $\alpha_i - (\sigma - 1)$). Thus, the overall response – the city-level trade elasticity – is equal to α_i , i.e., the shape parameter of the local productivity distribution. Since α_i also drives differences in city size (with larger cities having lower α_i), the trade elasticity is *decreasing* with city size. Intuitively, the thicker upper tail in larger cities implies a relatively higher mass of firms above the export cut-off. Thus, the mass of firms that enter the export market following a reduction in variable trade costs is lower relative to the mass of firms that are already exporting. Since the intensive-margin-response is the same across cities, this smaller extensive-margin-response implies that the overall trade elasticity is lower in larger cities.

The country-level trade elasticity can be decomposed as the trade-weighted average of city-level trade elasticities. Since these city-level elasticities are heterogeneous across cities and the distribution of population and economic activity across cities depends on international trade costs, the country-level trade elasticity is not constant and depends on the overall level of international trade costs.

4.4 Model Validation: City Size and Trade Elasticity

We can use Proposition 5 to validate our model, as it provides a novel prediction that arises naturally from our key assumptions. At the same time, the trade elasticity does not vary by city size in workhorse spatial models without differences in the upper tail of the firm productivity distribution across locations (e.g., Allen and Arkolakis, 2014; Redding, 2016).

To empirically test whether the trade elasticity is lower in larger cities, we need exogenous variation in the trade costs of exporters. We exploit the granting of permanent normal trade relations

(PNTR) for Chinese exports to the US in 2001. As pointed out by [Pierce and Schott \(2016\)](#), the conferral of PNTR removed uncertainty regarding future increases in US duties on imports from China and thereby promoted Chinese exports. Building on the identification strategy in [Pierce and Schott \(2016\)](#), we run the following regression:

$$\ln(\text{exports}_{cjt}) = \alpha [PostPNTR_t \times NTRGap_j] + \beta [PostPNTR_t \times NTRGap_j \times \ln(pop_c)] + \delta \mathbf{X}_{cjt} + \gamma_j + \gamma_t + [\gamma_c] + \varepsilon_{cjt} \quad (22)$$

where the dependent variable $\ln(\text{exports}_{cjt})$ denotes the exports of industry j in city c in year t , and pop_c is city population in the year 2000. $PostPNTR_t$ is a dummy that takes on value one for $t \geq 2001$, and $NTRGap_j$ is the difference (for industry j) between the US most favored nation (MFN) tariffs and non-MFN tariff rates. The former became permanent under PNTR. Thus, $NTRGap_j$ measures the counterfactual increase in trade costs that would have occurred if China's normal trade relations (NTR) status had not been extended. We calculate this NTR gap at the 3-digit Chinese industry level using data from [Pierce and Schott \(2016\)](#).³⁷ Non-MFN tariffs were on average 34% across the 112 Chinese industries in 1999, while MFN tariffs were merely 4%. Importantly for our identification, $NTRGap_j$ exhibits substantial variation across industries, with mean and standard deviation of 30% and 14%, respectively. We standardize $NTRGap_j$ and $\ln(pop_c)$ for a straightforward interpretation of coefficients.

We run regression (22) for the years 1998, 2000, 2003, 2006, and 2007. The start year is determined by data availability. We stop in 2007 in order to avoid the negative export demand shock induced by the Great Recession. The vector \mathbf{X}_{cjt} includes city-level controls (city population, distance to the coast, distance to the border, and capital-to-labor ratio, all in logs), as well as each of these variables' interactions with $PostPNTR_t$, $NTRGap_j$, and with $PostPNTR_t \times NTRGap_j$ (which is the reason why \mathbf{X}_{cjt} has subject cjt). Finally, γ_j , γ_t , and γ_c are industry, year, and city fixed effects, respectively. We report both regressions with and without city fixed effects. Note that these absorb the city-level controls but not their interactions with $PostPNTR_t$ and $NTRGap_j$. We weigh each city-industry cell by its share in overall exports of the industry, and we cluster standard errors at the industry as well as at the city level.

Holding everything else equal, we expect a larger response of exports post-PNTR in sectors with higher $NTRGap_j$ – that is, we expect a positive coefficient α . However, our main coefficient of interest in testing Proposition 5 is β , which measures how the response of city-industry-level exports to PNTR varies with city size (pop_c). We expect a negative coefficient ($\beta < 0$), corresponding to a declining trade elasticity with city size. Table 6 presents our results. We begin with a parsimonious specification in column 1 that controls only for industry and year fixed effects, and then

³⁷We match the original HS6 products to 3-digit Chinese industries using the crosswalk from [Dean and Lovely \(2010\)](#). Overall, this yields NTR gap data for 112 industries.

introduce city-level controls \mathbf{X}_{cjt} (col 2), city fixed effects (col 3), and add the interactions of city controls with $PostPNTR_t$ and $NTRGap_j$ (cols 4 and 5). The main results are similar throughout. In order to interpret coefficients, we standardize the $NTRGap_j$ variable and $\ln(pop_c)$. According to the estimate for α in Table 6, a one-standard-deviation (std) increase in $NTRGap_j$ led to a 13-18 percent increase in exports post-PNTR (for a city with average population size). This serves as a validation of our approach, as the reduction in uncertainty about U.S. import tariffs indeed led to higher exports from China. Most importantly, this effect is weaker in larger cities, as shown by the statistically significant, negative coefficients on the interaction (β): For small cities (in the bottom decile of population), the net effect of a one-std higher $NTRGap_j$ is a 44% increase in exports post-PNTR, while it is only 8% for large cities (in the top decile). Thus, the reduction in trade elasticities in larger cities is not only statistically but also economically significant. Overall, the results in Table 6 strongly support the prediction of Proposition 5. This validates our core mechanism of cross-city differences in the tail of firm productivity distributions – the distinguishing feature of our model that gives rise to Proposition 5.

5 Quantitative Analysis

In this section, we embed the mechanism we highlighted in our simple model in an open-economy quantitative spatial model, which we then fit to Chinese and French microdata. We evaluate its' key predictions quantitatively, studying the relationship between city size and i) export intensity and ii) the local trade elasticity. We also study the effects of counterfactual trade and spatial policies on welfare, productivity, wage inequality, spatial concentration and the aggregate trade elasticity. Finally, we compare our results to more standard models studied in spatial economics and international trade.

5.1 Quantitative Model

In order to make our framework amenable to quantitative analysis, we adapt the simple model presented in the previous section in four ways: First, we model the export market as one potentially asymmetric country (subscript c) without internal geography, while the home country consists of N locations. Second, we allow for a full matrix of bilateral domestic iceberg trade costs across the N locations: d_{ni} units of a commodity need to be shipped from origin i for one unit of the commodity to arrive at destination n . We assume exporting happens through specific locations that are located at the border or feature large ports in the data. Specifically, the cost of exporting from location i is equal to $d_{ci} = d_{pi} \times \tau$, where d_{pi} is the distance to the closest point of exit and τ is the cost of exporting. This allows the model to reflect that a location might be more export-intensive because it is located close to a port. Third, we allow for a richer specification of productivity differences across locations. While we maintain the assumption that firm productivity is drawn from location-specific Pareto distributions, we allow these distributions to differ across

locations not only in their shape parameters (α_i), but also in their scale parameters (A_i). A higher *scale* parameter A_i increases average city productivity while having no effect on export intensity. As we explained in our simple model, export intensity is solely driven by the *shape* parameter α_i , which *ceteris paribus* not only renders a thicker upper tail but also higher average productivity. The quantitative model effectively allows A_i to counteract this hard-wired effect of α_i on average productivity. Fourth, we introduce amenities in the model that vary across locations, and we allow workers to have heterogeneous preferences for different locations, which renders local labor supply imperfectly elastic and allows for differences across locations in real wage as well as ex-post utility.

Preferences. Worker preferences are given by:

$$U_i(\omega) = b_i(\omega)c_i(\omega)^\beta h_i(\omega)^{1-\beta}, \quad (23)$$

with c_i and h_i as defined above, and $b_i(\omega)$ denoting location-worker-specific amenity shocks that are drawn independently from a Fréchet distribution:

$$F_i(b) = e^{-B_i b^{-\epsilon}} \quad (24)$$

The scale parameters B_i determine average amenity levels for each location i , and the shape parameter ϵ controls the dispersion of amenities across workers for each location. With this specification of preferences, the indirect utility function at each location is given by:

$$U_i(\omega) = \frac{b_i(\omega)v_i}{P_i^\beta (p_i^h)^{1-\beta}}, \quad (25)$$

where P_i denotes the price index, which is now location-specific due to within-country trade costs. As in our simple model, v_i denotes total income per worker, and p_i^h is the price of housing at location i .

Given the extreme-value assumption on the amenity draws, utility is distributed Fréchet. Within the home country, the population share of each location is given by

$$\frac{L_i}{L} = \frac{B_i(v_i/P_i^\beta (p_i^h)^{1-\beta})^\epsilon}{\sum_{k \in N} B_k(v_k/P_k^\beta (p_k^h)^{1-\beta})^\epsilon} \quad (26)$$

The expected utility at each location is given by

$$\bar{U} = \delta \left[\sum_{k \in N} B_k(v_k/P_k^\beta (p_k^h)^{1-\beta})^\epsilon \right]^{\frac{1}{\epsilon}} \quad (27)$$

where $\delta \equiv \Gamma((\epsilon - 1)/\epsilon)$, with $\Gamma(\cdot)$ denoting the Gamma function.

General equilibrium of the quantitative model. The general equilibrium is represented by a measure of workers for each location L_i together with the following sets for each location: domestic market entry thresholds for firms ψ_i^{d*} , entry-into-exporting thresholds ψ_i^{x*} , firm entrants M_i^e , wages w_i , price indices P_i , and land rents r_i , as well as their corresponding variables for the rest of the world $w_c, \psi_c^{d*}, \psi_c^{x*}, M_c^e, P_c, r_c$. The equilibrium is determined by the following conditions in each location: trade balance, zero profit condition for serving the local and the export market, firm free entry condition, price index equation, land market clearing, location choice probabilities, and labor market clearing (see D.1 for the system of equations).

Welfare gains from trade in the quantitative model. Our model delivers a simple formula for welfare changes from changing trade costs:

$$\frac{\bar{U}^T}{\bar{U}^{T'}} = \left(\frac{\pi_{ii}^{T'}}{\pi_{ii}^T} \right)^{\frac{\beta}{\alpha_i}} \left(\frac{L_i^{T'}}{L_i^T} \right)^{\frac{1}{\epsilon} + (1-\beta) - \frac{\beta}{\sigma-1}} \Omega_i \quad (28)$$

where we have shut down the housing construction channel (i.e., set $\gamma = 1$) to facilitate comparison across models. T and T' refer to two equilibria with different levels of trade costs. Ω_i is an artifact of the assumption that there are no additional fixed costs of entering the market in other domestic locations.³⁸ The welfare changes in (28) is closely related to similar formulas from the literature: Relative to a trade model without geography (e.g., Melitz, 2003; Chaney, 2008), welfare changes do not only depend on changes in the own-trade share (π_{nn} , which in our setting is defined locally rather than nationally), but also on changes in the distribution of population in space (L_n). Relative to standard economic geography models, (e.g. Redding, 2016), the gains from trade in our setting also depend on the heterogeneous trade elasticity across location, given by the Pareto shape parameter α_i . We show below that these differences matter quantitatively below, underlining the importance to account for our mechanism.

5.2 Calibration

Pareto shape parameter. A central object for our quantitative analysis is the value for the Pareto shape parameter in each location, α_i . To estimate these parameters, we follow Head, Mayer, and Thoenig (2014), estimating QQ regressions for each location using firm-level domestic revenue data and accounting for industry fixed effects. We then multiply the resulting coefficient by $(\sigma - 1)$

³⁸This term is given by $\Omega_i = \left[\frac{R_i^{T'}(P_i^{T'})^{\sigma-1}}{R_i^T(P_i^T)^{\sigma-1}} \frac{\sum_k R_k^T(P_k^T)^{\sigma-1} d_{ki}^{1-\sigma}}{\sum_k R_k^{T'}(P_k^{T'})^{\sigma-1} d_{ki}^{1-\sigma}} \right]^{\beta \left(\frac{1}{\sigma-1} - \frac{1}{\alpha_i} \right)}$. If we introduced a fixed cost of entering the domestic market in location i , f_i , then we would obtain $\Omega_i = 1$. We refrain from this assumption with our quantitative exercise in mind, as city-specific fixed entry costs would be difficult to pin down empirically. In the following model comparisons, we abstract from Ω_i , as it reflects second-order effects.

following di Giovanni, Levchenko, and Mejean (2014) to recover the underlying productivity distribution Pareto shape parameter. In a nutshell, the QQ method estimates the domestic sales Pareto distribution parameters, minimizing the distance between the empirical and theoretical quantiles.

Table A.9 shows the summary statistics of the estimated Pareto shape parameters for China and France. We find significant dispersion of the estimates across cities in both countries, with a mean of 3.24 for China (2.95 for France) and a standard deviation of 0.68 (0.36 for France). Importantly, as Figure 2 illustrates, the estimated Pareto shape parameters show a strong negative, statistically highly significant correlation with city size in both countries, with estimated coefficients at -0.17 for China and -0.20 for France. In other words, large cities feature productivity distributions with thicker upper tails (i.e., smaller shape parameters). Note that these estimates stem directly from the data, without imposing any structure from our model. Thus, the strong negative correlation in Figure 2 directly supports our central assumption.

Other parameters. The rest of the calibration extends Redding (2016) for the case of heterogeneous firms with fixed production and export costs. We relegate the technical details to Appendix D.3. In short, the procedure inverts the model using the values of externally calibrated parameters together with data on population (L_i), wages (w_i), and country-level exports to recover the locational fundamentals: amenities (B_i), the Pareto *scale* parameter (A_i), and the value for the variable trade cost between ports and the rest of the world, τ .³⁹

5.3 Fitting the Stylized Fact: Export Intensity and City Size

In this section, we evaluate the fit of the model to our key stylized fact: the higher export intensity of larger cities. Note that the relationship between export intensity and city size was not a targeted moment in our calibration exercise. Thus, matching this empirical moment is a stringent test for our model and its underlying mechanisms. Figure 3 plots binned scatter diagrams for city-level export intensity and city size for China and France, respectively. Each dot in the graph corresponds to approximately 10 cities in China and 5 cities in France.⁴⁰ The model produces a good fit for the relationship between export intensity and city size. For both China and France, the model can account for more than two-thirds of the correlations observed in the data. The elasticities of export intensity with respect to city size for China and France, respectively, are 0.202 and 0.167 in the model, as compared to 0.272 and 0.185 in the data.

To further evaluate the performance of our model, we now turn to model-based extensive-intensive margin decompositions similar to the one presented in Section 3.4. We turn off specific

³⁹Table A.8 summarizes the values of the main parameters and the corresponding data moments. We display the correlation of the resulting fundamentals (A_i, B_i) with city size in Figure A.5: As we would expect, the level of amenities and the productivity *scale* parameter correlate positively with city size.

⁴⁰To make model and data directly comparable in the figure, we (i) normalize the average export intensity across cities to zero both in the data and in the model (our model matches the aggregate sales-weighted export intensity, leading to slight differences in the average level as compared to the data), and (ii) we control for the same set of geographic characteristics included in Table 2 when plotting the data.

features in our quantitative model one by one and examine how that affects its performance relative to the data. This allows us to shed light on the role played by the various mechanisms in matching the empirical relationship between export intensity and city size. Table 7 presents the correlation between export intensity and city size in the data (row 1 in each panel) and in the different versions of our model (rows 2-5). We report results including the geographical controls used in the empirical analysis (but the results are very similar when excluding these controls). We begin with our baseline model (row 2). For both China and France, most of the correlation between export intensity and city size is driven by the importance of exporters (extensive margin) rather than the importance of exporting for exporters (intensive margin). This broadly reflects the patterns in the data: For France, the model matches both margins very well; for China, there is a negative correlation between city size and the intensive margin in the data, while the model predicts a modest positive relationship.

Next, we compare the fit of our baseline model to the fit of two polar opposite benchmarks. The first benchmark (“Homogeneous firms model”) is a model where firms are *homogeneous* within cities, but there can be variation in productivity levels across cities.⁴¹ This benchmark is similar to the increasing returns model in Redding (2016), allowing us to check whether differences in foreign- relative to domestic-market access can produce the observed relationship between export intensity and city size. The results (row 3 of Table 7) show that without firm heterogeneity, the model can only account for a small fraction of our stylized fact: at most 8% (0.022/0.272) in China and less than 4% in France (0.007/0.185). In addition, the small correlation between export intensity and city size is entirely driven by the intensive margin. This was to be expected because in the absence of within-city firm heterogeneity, all firms engage in export markets in the same way, and there is no selection-into-exporting (extensive-margin) mechanism. These results suggest that differential market access can, at most, play a very modest role in accounting for the observed relationship between export intensity and city size.

Our second benchmark (“Baseline without internal trade costs”) is a model in which the role of market access is neutered by assuming costless within-country trade, while city-level productivity distributions are kept the same as in our baseline model. The results in row 4 of Table 7 show that the remaining core mechanism – heterogeneity in firm-level productivity distributions across cities – can account for up to 50 percent of the elasticity between export intensity and city size in China, and for about 90 percent in France.⁴² Finally, we explore the role of cross-city heterogeneity in the thickness of the upper tail of the firm productivity distributions – the key component in our

⁴¹We use the same calibration strategy as in the baseline model. Since the model has no firm heterogeneity, we do not need to fit the shape parameter but set it equal to infinity, which implies a degenerate productivity distribution for all cities. We keep the structure of the internal variable trade costs between domestic locations the same as in our baseline model.

⁴²Note that in this model, all effects operate via the extensive margin, as symmetric market access produces a null intensive margin: conditional on exporting, firms at all locations export the same fraction of their production.

theory. We introduce a third benchmark (“Baseline with common shape parameter”): a model featuring heterogeneity in market access across cities, within-city firm-level heterogeneity, but *no* differences in the thickness of the upper tail (we allow *scale* parameters to differ across cities but impose a common *shape* parameter). As shown in row 5 of Table 7, this setup can explain only a small fraction of the correlation between city size and export intensity. In addition, this is almost entirely driven by the intensive margin of trade, thus contradicting the pattern in the data. This result underlines that differences in the productivity upper tail are crucial for our stylized fact.

Taken together, our results suggest that allowing firm productivity distributions to vary across cities – and in particular allowing cities to have differences in the thickness of the upper tail of their firm productivity distributions – is key to accounting for the spatial concentration of export activity in larger cities. This novel mechanism, in isolation, allows us to account for most of the reduced-form relationship between export intensity and city size, while differences in foreign- relative to domestic-market access play only a minor role.

5.4 Validating the Model: City Size and Local Trade Elasticity

Under the assumptions made in Section 4, the model predicts a decreasing relationship between city size and the local trade *elasticity*, i.e., exports in larger cities are *less responsive* to changes in trade costs. To test whether this prediction holds under the more general assumptions of our quantitative model and to assess its quantitative importance, we calculate the trade elasticity for each city from a marginal decrease in trade costs at the observed equilibrium. These estimates provide another untargted moment that allows us to confront our quantitative model with the PNTR regression results discussed in Section 4.4. Figure 4 plots the local trade elasticity from our full quantitative model against city size. For China, the estimated gradient is 0.118 (0.156 for France). This is very similar to our PNTR regressions, where we find a gradient of 0.11.⁴³

Finally, in Appendix D.5 we show that the relationship between city size and trade elasticities is a distinguishing feature of our model, as compared to the other models discussed in Section 5.3: The city-size-elasticity relationship only results from our baseline quantitative model (even in a simpler version without internal trade costs). Crucially, the relationship becomes minuscule (and reverses its sign) when we use a common Pareto shape parameter across cities, or when using an homogeneous-firms model.

5.5 Counterfactual Analysis

This section analyzes the effects of commonly discussed trade and spatial policies in the context of our model. We illustrate how the aggregate productivity and welfare implications of these policies

⁴³The city-size-gradient coefficients (i.e., the triple interactions) in Table 6 are negative, reflecting a smaller reaction of exports to *falling* trade costs (*NTRGap*) in larger cities (i.e., a smaller trade elasticity in absolute terms). In the model, the size- gradient-coefficients are positive, reflecting a less negative trade elasticity (i.e., also smaller in absolute terms).

differ relative to the predictions of existing workhorse models.

Trade Liberalization

We first study the effects of trade openness by comparing the actual trade equilibrium to an autarky counterfactual. This counterfactual is identical to the actual equilibrium along all key parameters, but international trade costs are assumed to be prohibitive. To contrast our model’s counterfactual predictions, we compare the quantitative effects of trade in this model to four benchmarks commonly considered in the literature. We perform these exercises through the lens of a researcher fitting the different model alternatives to the data (that is, we separately calibrate each model to the data; see Appendix E for details). Table 8 shows the results for our baseline model and four benchmarks commonly considered in the trade and spatial economics literature. (2) A model with the same rich internal geography as in our baseline but with homogeneous firms.⁴⁴ (3) A model equal to the baseline model except that it does not feature any domestic trade costs and, therefore, no differences in the cost of exporting across locations. (4) A heterogeneous firm model with geography that only features differences in the level of productivity but not in the tail of the productivity distribution across cities (i.e., we restrict the shape parameter of the Pareto distribution to be constant across locations while the location parameter is location-specific). (5) A model featuring heterogeneous firms but with no internal geography (i.e. the entire country is modeled as a single location), which therefore closely resembles traditional trade models in the spirit of Melitz (2003) and Chaney (2008).

We first discuss the welfare gains from trade (column 1 in Table 8). Our baseline model predicts gains equal to 4.4% and 4.3% for China and France, respectively. In all alternative models, the welfare gains from trade are meaningfully larger than in our model, especially in the case of China. Gains from trade arise as a reduction in trade cost increases the value of exports and imports in the economy. Since in our model (as in the data), exporters are spatially concentrated in larger cities, exporting firms face relatively high local wages due to high housing costs, and those increase overproportionally in large cities when exports expand following a reduction in trade costs. In other words, our model implies that the spatial concentration of exporting in large cities limits the ability to expand exporting, which dampens the gains from trade. This highlights that our mechanism has important implications for aggregate statistics.

Opening up to trade reallocates workers to larger cities and increases the nominal wage gap as well as the real income gap between smaller and larger cities. To quantitatively evaluate this reallocation, we report the effect of trade on the ratio between the 90th and the 10th percentile of city-level population, nominal wages, and real wages in columns 2-4 in Table 8.⁴⁵ In our baseline

⁴⁴Conceptually, this benchmark is closely related to the increasing-returns-to-scale model presented in Redding (2016). We choose the increasing returns model as it is closest to our model.

⁴⁵Figure A.7 visualizes the differences in population reallocation from autarky to trade along the city size distribu-

model, trade increases the population ratio by 10.0% in China and by 6.9% in France, while the nominal wage ratio increases by 1.9% and 2.3%. The real income ratio increases by 2.4% in China and 2.2% in France. Note that the difference between nominal and real wage changes is not just driven by house prices (which rise in larger relative to smaller cities), but also due to relative changes in the price index of the tradable good, whose change is not monotonically tied to city size. These results highlight the important effect that international trade has on regional inequality. Following a reduction in trade cost, smaller cities shrink while larger cities grow, and workers who remain in smaller cities experience a reduction in their real income while workers in larger cities experience real wage gains. All competing models deliver smaller spatial reallocation and changes in nominal wages / real income due to trade liberalization. Interestingly, restricting the upper tail of the productivity distributions to be the same across cities (row 4) leads to similar effects as a model with homogeneous firms (row 2).⁴⁶ Shutting down internal trade costs (row 3) also reduces spatial reallocation, although significantly more in China than in France.

In our model, changing trade costs affect the aggregate trade elasticity. Figure 5 illustrates this prediction, plotting the country-level trade elasticity and the spatial concentration of population for different levels of trade costs (τ_c). As trade costs decline (moving to the left on the x-axis), economic activity and exporting become more concentrated in larger cities. As we have shown both theoretically (in Proposition 5) and empirically (in our PNTR regression in Table 6), the trade elasticity is smaller in larger cities. Thus, when overall trade costs decline, the aggregate trade elasticity falls in absolute terms. This result complements recent findings highlighting how the trade elasticity changes with trade costs (Adão et al., 2020), providing a novel, spatial mechanism.

Spatial Policies

Our second counterfactual exercise studies the effects of reducing land-use restrictions on aggregate productivity and export intensity at the country level. We implement this policy as a reduction in the parameter γ , which measures the intensity of land use in the housing production function, in the home country. Reducing γ tilts the housing supply curves such that the price of housing increases less steeply with population. As a result, the cost of living and, therefore, wages increase less steeply with city size, as workers require lower wages to be willing to move to larger cities. This, in turn, increases labor demand from firms in larger cities relative to smaller cities, leading to a reallocation of population to the former. Overall, this affects aggregate export intensity in a number of ways: First, since firms in larger cities are more productive and are more likely to export, this pushes up country-level export intensity (this is the only mechanism that is also active in our simple model). Second, this reallocation increases total income in the domestic country and

tion.

⁴⁶Intuitively, the firm heterogeneity in our model by itself does not lead to a heterogeneous transmission of an aggregate trade shock. However, it amplifies the effects induced by differences in variable trade costs through the selection margin.

thereby raises domestic relative to foreign demand, reducing aggregate export intensity. Third, this positive effect on aggregate export intensity gets attenuated as the additional entry into the export market reduces the price index in the foreign country and thereby the exports of all firms. Finally, the spatial reallocation of factors also affects domestic sales across locations in an ex-ante ambiguous way. Thus, while total exports unambiguously increase, the overall effect of planning deregulation on export *intensity* is ambiguous.

To quantitatively explore these effects, we increase the housing supply elasticity in the model from the 25th to 75th percentile of previous estimates. For China, we use the range of elasticities estimated by Wang, Chan, and Xu (2012). For France, we use estimates based on US data from Saiz (2010). The results are displayed in Table 9. Column 1 shows that the concentration of population in large cities increases substantially. Next, columns 2-4 show that for China, aggregate exports grow more than sales, so that export intensity increases as a result of deregulated urban planning. For France, in contrast, sales grow more than exports, so that export intensity declines. Finally, aggregate productivity (column 5) increases particularly strongly (by 6%) in China as a result of the more elastic housing supply. This relatively large number (as compared to 1.6% in France) is likely not only a result of housing supply constraints, but also of China’s restrictive Hukou system, which limits migration to productive cities. These aggregate effects are towards the lower end of findings in the existing literature, largely focused on the US. In a model with endogenous housing supply restrictions, Parkhomenko (2023) finds that lowering regulations in a group of 10 US ‘superstar’ cities would increase productivity by 3.6%. By contrast, Hsieh and Moretti (2019) find that reducing regulations in New York, San Francisco and San Jose alone would raise output by 3.7%-8.9% while Herkenhoff, Ohanian, and Prescott (2018), in a state level analysis, find that reducing regulations everywhere in the US to half the level of Texas would increase aggregate productivity by 12.4% to 19.4%.

6 Conclusion

Trade policy has received renewed interest in recent years, as globalization has been blamed for widening disparities in many developed countries (Autor et al., 2013; Ezcurra and Pose, 2013; Dix-Carneiro and Kovak, 2017a; Potlogea, 2018). In response to this interest, a nascent literature has begun to analyze the interplay between trade and economic geography within countries. We contribute to this literature in three ways. First, using information from four major trading nations – China, France, the United States, and Brazil – we have documented a novel and highly robust stylized fact: Exporting is more unevenly distributed than overall economic activity, and in particular, it is disproportionately concentrated in larger cities. Importantly, this stylized fact is not driven by larger cities benefitting from better foreign market access. Second, we show that a relatively simple framework – an extension of the standard quantitative spatial equilibrium framework to in-

clude firm heterogeneity and a mechanism of selection into exporting in the spirit of Melitz (2003) – can explain this stylized fact. Third, we structurally estimate the model and use it to undertake counterfactual policy analyses.

Our model is designed to assess the effects of both trade policies and (domestic) spatial policies, giving rise to novel interactions between these two levers. We find that the corresponding welfare implications are richer and differ from those in the more parsimonious standard models that are nested in our framework: a standard trade model that ignores within-country geography, and an economic geography model that shuts down international trade.

Our theoretical framework opens the door for future work that exploits the interplay of international trade and domestic economic geography. For example, our model naturally lends itself to exploring the rich interactions between policies that reduce internal trade costs (e.g., infrastructure investments) and those that affect international trade costs (e.g, trade agreements).

References

- Adão, R., C. Arkolakis, and S. Ganapati (2020). Aggregate Implications of Firm Heterogeneity: A Non-parametric Analysis of Monopolistic Competition in Trade Models. Technical report, National Bureau of Economic Research.
- Allen, T. and C. Arkolakis (2014). Trade and the Topography of the Spatial Economy. *The Quarterly Journal of Economics* 129(3), 1085–1140.
- Anderson, J. E. and E. van Wincoop (2004). Trade Costs. *Journal of Economic Literature* 42(3), 691–751.
- Au, C.-C. and J. V. Henderson (2006). Are Chinese Cities Too Small? *Review of Economic Studies* 73(3), 549–576.
- Autor, D. H., D. Dorn, and G. H. Hanson (2013). The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review* 103(6), 2121–2168.
- Bai, Y. and R. Jia (2021). The Economic Consequences of Political Hierarchy: Evidence from Regime Changes in China, AD1000-2000. Technical report. forthcoming.
- Bakker, J. D. (2020). Trade and Agglomeration: Theory and Evidence from France.
- Behrens, K., G. Duranton, and F. Robert-Nicoud (2014). Productive Cities: Sorting, Selection, and Agglomeration. *Journal of Political Economy* 122(3), 507–553.
- Behrens, K., C. Gaigne, G. I. Ottaviano, and J.-F. Thisse (2006a). How density economies in international transportation link the internal geography of trading partners. *Journal of Urban Economics* 60(2), 248–263.
- Behrens, K., C. Gaigne, G. I. Ottaviano, and J.-F. Thisse (2006b). Is remoteness a locational disadvantage? *Journal of Economic Geography* 6(3), 347–368.
- Behrens, K., C. Gaigne, G. I. Ottaviano, and J.-F. Thisse (2007). Countries, regions and trade: On the welfare impacts of economic integration. *European Economic Review* 51(5), 1277–1301.

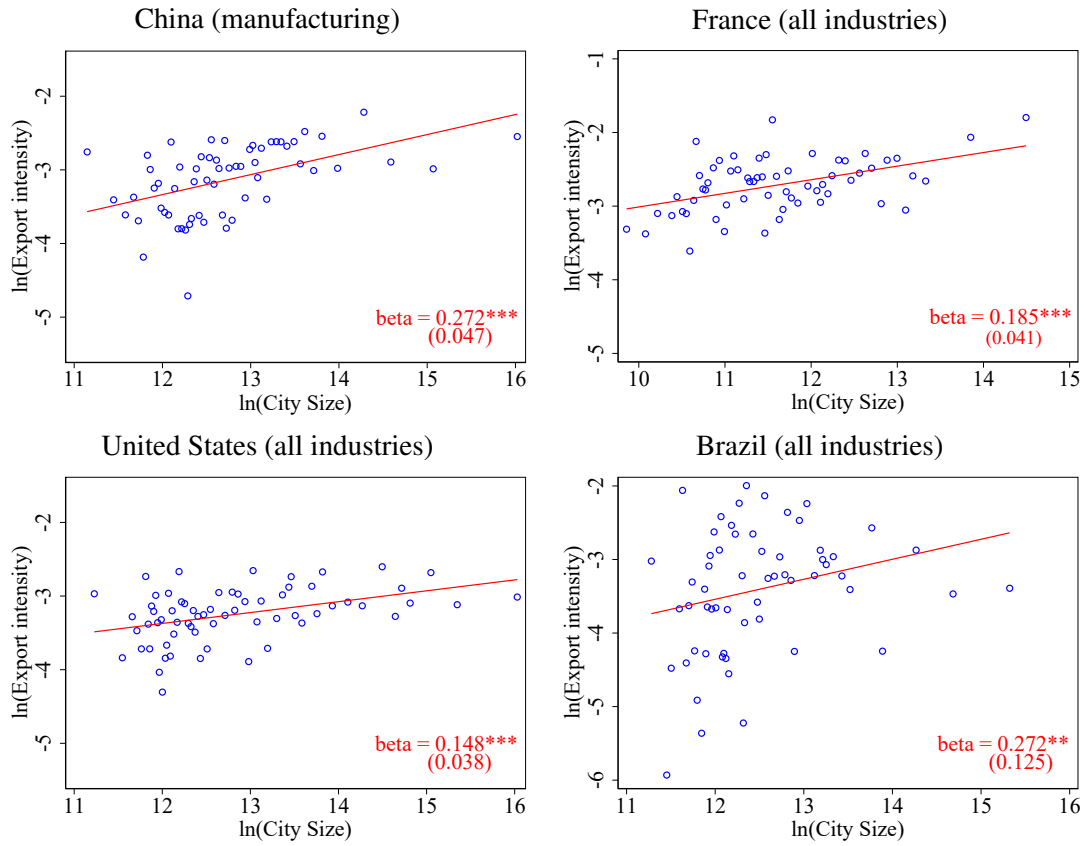
- Behrens, K., C. Gaigne, G. I. Ottaviano, and J.-F. Thisse (2009). Industry location and welfare when transport costs are endogenous. *Journal of Urban Economics* 65(2), 195–208.
- Bellemare, M. F. and C. J. Wichman (2020, February). Elasticities and the inverse hyperbolic sine transformation. *Oxford Bulletin of Economics and Statistics* 82(1), 50–61.
- Bernard, A. B. and J. B. Jensen (1999). Exceptional Exporter Performance: Cause, Effect, or Both? *Journal of International Economics* 47(1), 1–25.
- Brandt, L., J. Van Biesebroeck, and Y. Zhang (2014). Challenges of working with the Chinese NBS firm-level data. *China Economic Review* 30(C), 339–352.
- Bryan, G. and M. Morten (2019). The Aggregate Productivity Effects of Internal Migration: Evidence from Indonesia. *Journal of Political Economy* 127(5), 2229–2268.
- Chaney, T. (2008). Distorted Gravity: The Intensive and Extensive Margins of International Trade. *American Economic Review* 98(4), 1707–1721.
- Chauvin, J. P., E. Glaeser, Y. Ma, and K. Tobio (2017). What is different about urbanization in rich and poor countries? Cities in Brazil, China, India and the United States. *Journal of Urban Economics* 98(C), 17–49.
- Chen, N., D. Novy, C. Perroni, and H. C. Wong (2023). Urban-biased structural change.
- Chen, Z. (2024). Economic growth and the rise of large firms. *Unpublished manuscript*.
- Cheng, W. and A. Potlogea (2020). Trade Liberalization and Economic Development: Evidence from China’s WTO Accession.
- Combes, P.-P., G. Duranton, L. Gobillon, D. Puga, and S. Roux (2012). The Productivity Advantages of Large Cities: Distinguishing Agglomeration From Firm Selection. *Econometrica* 80(6), 2543–2594.
- Coşar, A. K. and P. D. Fajgelbaum (2016). Internal geography, international trade, and regional specialization. *American Economic Journal: Microeconomics* 8(1), 24–56.
- Costa, F., J. Garred, and J. P. Pessoa (2016). Winners and losers from a commodities-for-manufactures trade boom. *Journal of International Economics* 102(C), 50–69.
- Dauth, W., S. Findeisen, and J. Suedekum (2014). The rise of the east and far-east: German labor markets and trade integration. *Journal of the European Economic Association* 12(6), 1643–1675.
- Davis, D. R. and J. I. Dingel (2014). The Comparative Advantage of Cities. NBER Working Papers 20602, National Bureau of Economic Research, Inc.
- Davis, D. R. and J. I. Dingel (2019). A Spatial Knowledge Economy. *American Economic Review* 109(1), 153–170.
- Davis, M. A. and F. Ortalo-Magne (2011, April). Household Expenditures, Wages, Rents. *Review of Economic Dynamics* 14(2), 248–261.
- Dean, J. M. and M. E. Lovely (2010). Trade growth, production fragmentation, and china’s environment. In *China’s growing role in world trade*, pp. 429–469. University of Chicago Press.
- di Giovanni, J., A. A. Levchenko, and I. Mejean (2014). Firms, Destinations, and Aggregate Fluctuations. *Econometrica* 82(4), 1303–1340.
- Diamond, R. (2016). The determinants and welfare implications of us workers’ diverging location choices by skill: 1980–2000. *The American Economic Review* 106(3), 479–524.
- Dingel, J. I., A. Miscio, and D. R. Davis (2019). Cities, Lights, and Skills in Developing Economies. NBER Working Papers 25678, National Bureau of Economic Research, Inc.

- Dix-Carneiro, R. and B. Kovak (2017a). Trade Liberalization and Regional Dynamics. *American Economic Review* 107, 2908–2946.
- Dix-Carneiro, R. and B. K. Kovak (2015). Trade Liberalization and the Skill Premium: A Local Labor Markets Approach. *American Economic Review* 105(5), 551–557.
- Dix-Carneiro, R. and B. K. Kovak (2017b). Trade Liberalization and Regional Dynamics. *American Economic Review* 107(10), 2908–2946.
- Dix-Carneiro, R. and B. K. Kovak (2019). Margins of labor market adjustment to trade. *Journal of International Economics* 117(C), 125–142.
- Ducruet, C., R. Juhasz, D. K. Nagy, and C. Steinwender (2020). All aboard: The effects of port development.
- Eaton, J., S. Kortum, and F. Kramarz (2011). An Anatomy of International Trade: Evidence From French Firms. *Econometrica* 79(5), 1453–1498.
- Eckert, F., S. Ganapati, and C. Walsh (2022). Urban-biased growth: a macroeconomic analysis. Technical report, National Bureau of Economic Research.
- Eeckhout, J., R. Pinheiro, and K. Schmidheiny (2014). Spatial Sorting. *Journal of Political Economy* 122(3), 554–620.
- Ezcurra, R. and A. R. Pose (2013). Does Economic Globalization affect Regional Inequality? A Cross-country Analysis. *World Development* 52, 92–103.
- Fajgelbaum, P. and S. J. Redding (2022). Trade, structural transformation, and development: Evidence from argentina 1869–1914. *Journal of Political Economy* 130(5), 1249–1318.
- Garcia-Marin, A., A. V. Potlogea, N. Voigtländer, and Y. Yang (2020, December). Cities, Productivity, and Trade. NBER Working Papers 28309, National Bureau of Economic Research, Inc.
- Garcia-Marin, A. and N. Voigtländer (2019). Exporting and Plant-Level Efficiency Gains: It’s in the Measure. *Journal of Political Economy* 127(4), 1777–1825.
- Gaubert, C. (2018). Firm Sorting and Agglomeration. *American Economic Review* 108(11), 3117–3153.
- Giannone, E. (2022). Skilled-biased technical change and regional convergence. *CREI. Unpublished manuscript*.
- Head, K., T. Mayer, and M. Thoenig (2014). Welfare and trade without pareto. *American Economic Review* 104(5), 310–316.
- Helpman, E. (2018). *Globalization and Inequality*. Harvard University Press.
- Helpman, E. et al. (1995). *The size of regions*. Foerder Institute for Economic Research.
- Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating Trade flows: Trading Partners and Trading Volumes. *The Quarterly Journal of Economics* 123(2), 441–487.
- Henderson, J. V., D. Nigmatulina, and S. Kriticos (2021). Measuring urban economic density. *Journal of Urban Economics* 125(C).
- Herkenhoff, K., L. E. Ohanian, and E. C. Prescott (2018). Tarnishing the Golden and Empire States: Land-Use Regulations and the U.S. Economic Slowdown. *Journal of Monetary Economics* 93, 89–109.
- Hsieh, C.-T. and E. Moretti (2019). Housing Constraints and Spatial Misallocation. *American Economic Journal: Macroeconomics* 11(2), 1–39.
- Kovak, B. K. (2013). Regional Effects of Trade Reform: What Is the Correct Measure of Liberalization? *American Economic Review* 103(5), 1960–1976.

- Krugman, P. and R. Livas Elizondo (1996). Trade policy and the Third World metropolis. *Journal of Development Economics* 49(1), 137–150.
- Melitz, M. J. (2003). The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. *Econometrica* 71(6), 1695–1725.
- Melitz, M. J. and S. J. Redding (2015). New Trade Models, New Welfare Implications. *American Economic Review* 105(3), 1105–46.
- Monfort, P. and R. Nicolini (2000). Regional Convergence and International Integration. *Journal of Urban Economics* 48(2), 286–306.
- Moretti, E. (2013). Real wage inequality. *American Economic Journal: Applied Economics* 5(1), 65–103.
- Novy, D. (2013). International Trade without CES: Estimating Translog Gravity. *Journal of International Economics* 89(2), 271–282.
- Parkhomenko, A. (2023). Local causes and aggregate implications of land use regulation. *Unpublished manuscript*.
- Pavcnik, N. (2002). Trade Liberalization, Exit and Productivity Improvements: Evidence from Chilean Plants. *Review of Economic Studies* 69, 245–76.
- Pierce, J. R. and P. K. Schott (2016). The surprisingly swift decline of us manufacturing employment. *American Economic Review* 106(7), 1632–62.
- Potlogea, A. (2018). Globalization and the Skilled City. *Journal of Urban Economics* 107, 1–30.
- Redding, S. J. (2016). Goods trade, factor mobility and welfare. *Journal of International Economics* 101, 148–167.
- Redding, S. J. and D. M. Sturm (2008, December). The costs of remoteness: Evidence from german division and reunification. *American Economic Review* 98(5), 1766–97.
- Saiz, A. (2010). The Geographic Determinants of Housing Supply. *The Quarterly Journal of Economics* 125(3), 1253–1296.
- Schoefer, B. and O. Ziv (2022). Productivity, place, and plants. *Review of Economics and Statistics*, 1–46.
- Walsh, C. (2023). The startup multiplier.
- Wang, S., S. H. Chan, and B. Xu (2012). The estimation and determinants of the price elasticity of housing supply: Evidence from china. *Journal of Real Estate Research* 34(3), 311–344.

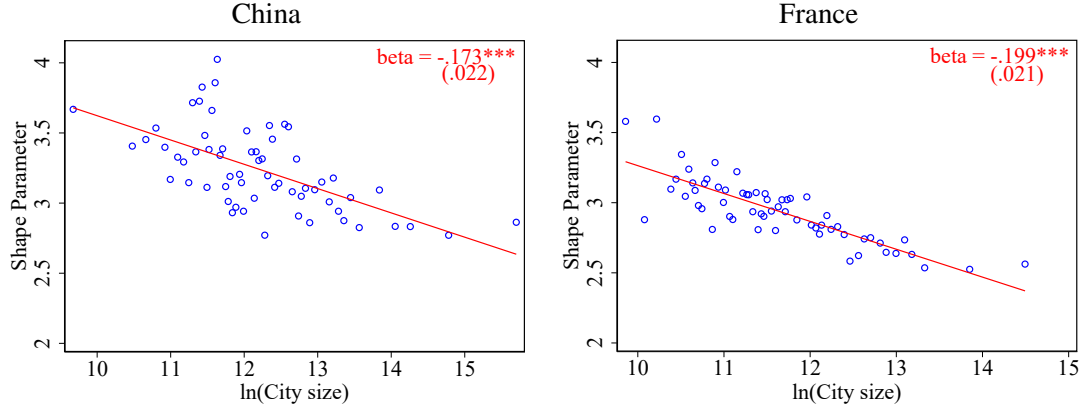
FIGURES

Figure 1: Export Intensity and City Size in China, France, the United States, and Brazil



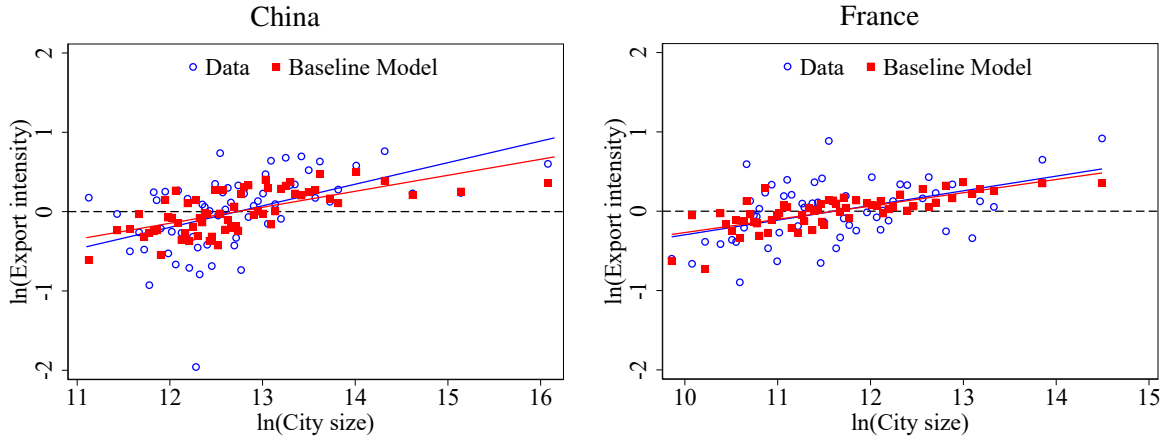
Notes: The figure shows a binned scatter plot between city-level ln export intensity (exports relative to revenues) and ln city size. Each dot (bin) represents 10 underlying cities for China, and 5 cities for the other three countries. Cities are defined in terms Microregions for Brazil, Metropolitan Areas for China and the United States (as defined by Dingel et al., 2019, using lights at night with a threshold equal to 30); and Commuting Zones for France. The analysis considers cities with positive exports and at least 250 manufacturing firms for China and France. For Brazil and the United States, the analysis considers cities with a population above 100,000 inhabitants. All figures include the following controls: ln average distance to other domestic cities, ln distance to the border, ln distance to the coast, border dummies, and coastal dummies.

Figure 2: Estimated Pareto Shape Parameter and City Size



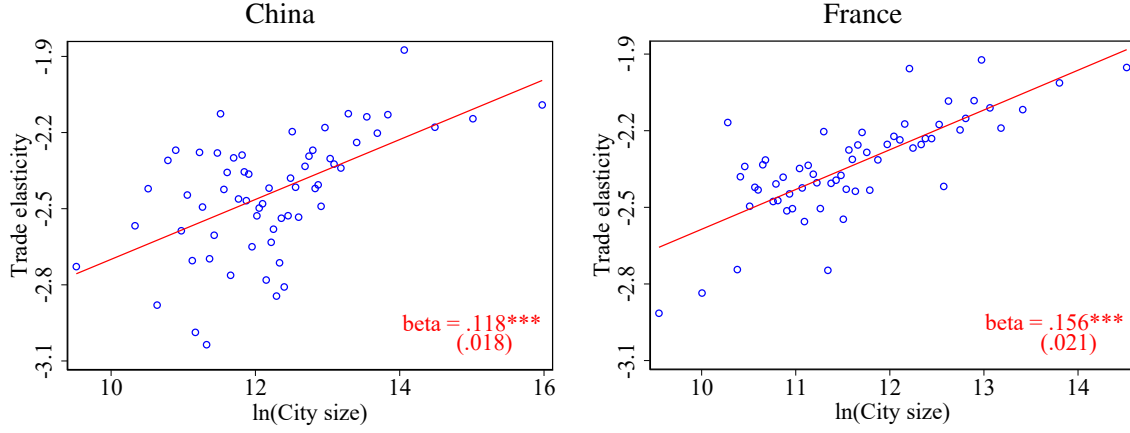
Notes: The figure provides evidence for a central assumption in our model: The fatter upper tail (i.e., lower Pareto shape parameter α_i) in larger cities. We show a binned scatter plot between the estimated α_i and city size for China and France. Each dot (bin) represents 10 underlying cities for China and 5 cities for France. See Section 5.2 for a description of the calibration procedure and Appendix D.3 for additional details.

Figure 3: Model vs. Data: (Predicted) Export Intensity and City Size



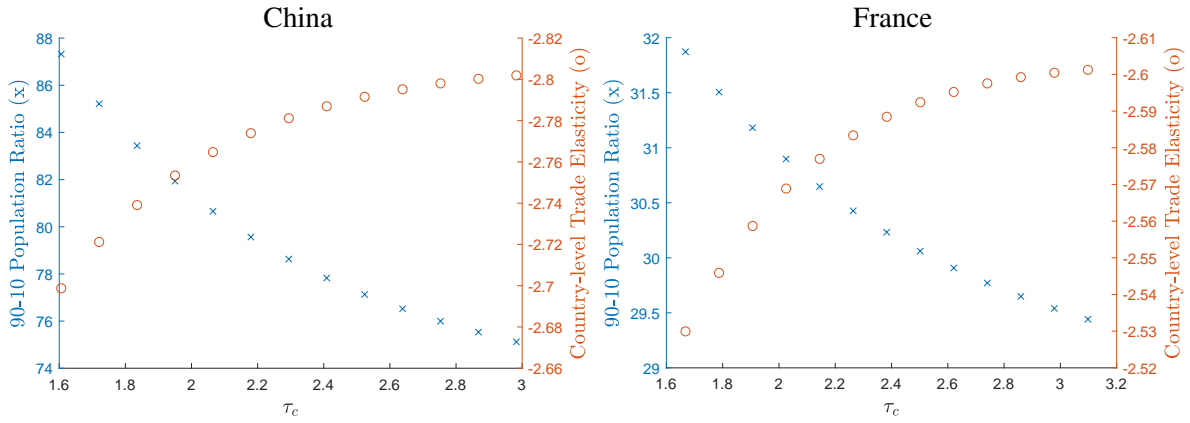
Notes: The figure plots a binned scatter plot between export intensity and city size for China and France for the empirical correlation in the data (blue-hollow circles) and the output of our quantitative model (red-solid squares), as presented in Section 5.1 and calibrated as described in Section 5.2. Each dot (bin) represents 10 underlying cities for China, and 5 cities for France. Both data and model output include linear regression line.

Figure 4: Model Results: Trade Elasticity vs. City Size for China and France



Notes: This figure plots the local international trade elasticity implied by our quantitative model (at the data-implied equilibrium) against the population size of each location for China and France. We calculate the trade elasticity for each city from a marginal decrease in trade costs at the observed equilibrium. Each dot (bin) represents 10 underlying cities for China, and 5 cities for France.

Figure 5: Effect of Trade Costs on Aggregate Trade Elasticity and Population Concentration



Notes: This figure illustrates the effect of international trade costs (τ_c – on the x-axis) on the aggregate (i.e., country-level) trade elasticity (right axis) and on the spatial concentration of economic activity (left axis – the population ratio of cities in the top decile relative to the bottom population decile).

TABLES

Table 1: City Size and Export Intensity: Descriptive Statistics

		Population size		Export intensity	
	(1)	(2)	(3)	(4)	(5)
	Obs.	Mean	Std. dev.	Mean	Std. dev.
China (Metropolitan areas)	629	709.3	2,732.4	8.59	9.75
France (Commuting zones)	304	192.5	374.5	8.40	6.13
United States (Metropolitan areas)	324	799.3	1,722.1	5.61	6.32
Brazil (Microregions)	316	463.6	1,129.1	8.99	12.72

Notes: The table shows statistics for the distribution of city sizes and export intensities in Brazil, China, France, and the United States. Population size is in '000s and export intensity in percent. Cities are defined in terms of Metropolitan Areas for China (as defined by [Dingel et al., 2019](#), using lights at night with a threshold equal to 30) and the United States, Commuting Zones for France, and Microregions for Brazil. The analysis considers cities with positive exports and at least 250 firms for China and France. For Brazil and the United States, the analysis considers cities with a population above 100,000 inhabitants.

Table 2: Export Intensity and City size in Brazil, China, France, and the United States

Dependent Variable: City-Level ln(Export Intensity)								
	— China —		— France —		— United States —		— Brazil —	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
ln City Size	.339*** (.052)	.272*** (.047)	.215*** (.043)	.185*** (.041)	.156*** (.037)	.148*** (.038)	.329*** (.121)	.272*** (.125)
Geography Controls		✓		✓		✓		✓
Mean Export Intensity:	.088	.088	.084	.084	.056	.056	.096	.096
R ²	.043	.264	.083	.184	.037	.155	.020	.208
Observations	615	615	304	304	324	324	296	296

Notes: The table analyzes the relationship between city size and export intensity, both in logs. See the note to Table 1 for the definition of cities in the four countries. City-level export intensity is defined as manufacturing exports over manufacturing sales for China; overall exports over sales for the France and the United States, and overall exports over (city-level) GDP for the case of Brazil. ‘Geography Controls’ include: ln(average distance to other domestic cities), ln(distance to the border), ln(distance to the coast), border dummies, and coastal dummies. Robust standard errors in parentheses. Key: ** significant at 1%; * 5%; * 10%.

Table 3: Export Intensity and City Size: 2SLS Results

Specification:	China				France			
	OLS (1)	RF (2)	FS (3)	IV (4)	OLS (5)	RF (6)	FS (7)	IV (8)
Dependent Variable	ln Export Intensity	ln Export Intensity	ln City Size	ln Export Intensity	ln Export Intensity	ln Export Intensity	ln City Size	ln Export Intensity
ln City Size	0.321*** (0.113)	—	—	0.361** (0.180)	0.148*** (0.041)	—	—	0.100* (0.052)
ln Historical Population	—	0.116** (0.0551)	0.321*** (0.0487)	—	—	0.097* (0.053)	0.968*** (0.052)	—
First Stage F-Statistic	—	—	43.4	—	—	—	348.5	—
Geog. Controls	✓	✓	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	-2.89	-2.89	14.27	-2.89	-2.43	-2.43	11.64	-2.43
R ²	.352	.324	.377	.352	.112	.091	.638	.148
Observations	260	260	260	260	297	297	297	297

Notes: This table examines the effect of city size on export intensity in China (columns 1 to 4) and France (columns 5 to 8), instrumenting contemporaneous city size with historical population. The analysis for China is run at the prefecture level, using 1580 prefecture-level population from [Bai and Jia \(2021\)](#) to instrument for city size. The analysis for France is run at the commuting zone level, using the population records for 1876 produced by INSEE as an instrument for current population. Cols 1 and 5 report OLS estimates. Cols 2 and 6 report the reduced form (RF), regressing ln export intensity on historical population. The first stage (FS) results of the Instrumental Variables (IV) regressions are reported in cols 3 and 7 together with the (cluster-robust) Kleibergen-Paap rKWD F-statistic. The corresponding Stock-Yogo value for 10% maximal IV bias is 16.4. ‘Geog. Controls’ are listed in the note to Table 2. The IV coefficients are reported in cols 4 and 8. Robust standard errors (in parentheses). Key: *** significant at 1%; ** 5%; * 10%.

Table 4: Decomposition into Within- and Between-Sector Export Intensity

Dep. Var.: Within- and Between Components of City-Level Export Intensity						
Specification:	China			France		
	(1) Within	(2) Between	(3) % Within	(4) Within	(5) Between	(6) % Within
ln(City Size)	.302*** (.0436)	-.031** (.0151)	>100%	.097*** (.023)	.088*** (.028)	52.4%
Geography Controls	✓	✓	—	✓	✓	—
Mean Dependent Variable	-0.95	-2.22	—	-0.29	-2.16	—
R ²	.214	.257	—	.171	.148	—
Observations	615	615	—	304	304	—

Notes: This table decomposes city-level export intensity into a within- and a between-industry component. Note that the coefficients on the two components add up to the overall coefficient in Table 2, col 2 (for China) and col 4 (for France). ‘Geography Controls’ are listed in the note to Table 2. Robust standard errors in parentheses. Key: *** significant at 1%; ** 5%; * 10%.

Table 5: Checking for the Role of Variable Transport Costs: Two Margins of Exporting

Dependent Variable as listed in table header						
	China			France		
	(1) $\ln\left[\frac{x_i}{r_i}\right]$	(2) $\ln\left[\frac{x_i}{r_i^x}\right]$	(3) $\ln\left[\frac{r_i^x}{r_i}\right]$	(4) $\ln\left[\frac{x_i}{r_i}\right]$	(5) $\ln\left[\frac{x_i}{r_i^x}\right]$	(6) $\ln\left[\frac{r_i^x}{r_i}\right]$
ln(City Size)	0.272*** (.0466)	-.179*** (.0331)	0.451*** (.0449)	.184*** (.041)	.018 (.0333)	.167*** (.0256)
Geog. Controls	✓	✓	✓	✓	✓	✓
Mean Dep. Var.	-3.14	-1.28	-1.85	-2.714	-1.69	-1.03
Shares	-1.28	-3.40	1.55	0.084	0.21	0.39
R ²	.12	.41	.11	.018	.05	.29
Observations	615	615	615	304	304	304

Notes: The table decomposes overall export intensity (x_i/r_i in cols 1,4 – with x_i and r_i denoting export revenues and total revenues, respectively, in city i) into the importance of exporting among exporters (the intensive margin x_i/r_i^x in cols 2,5 – with r_i^x denoting total (incl. domestic) revenues of exporters) and the remainder – the importance of exporters among local firms (r_i^x/r_i in cols 3,6). ‘Geog. Controls’ are listed in the note to Table 2. Robust standard errors are in parentheses. Key: ** significant at 1%; * 5%; * 10%.

Table 6: Empirical Test of Proposition 5: Heterogenous Local Trade Elasticity

Dependent Variable: $\ln(\text{exports}_{cjt})$					
	(1)	(2)	(3)	(4)	(5)
$\alpha : [PostPNTR_t \times NTRGap_j]$	0.18*** (0.051)	0.18*** (0.050)	0.16*** (0.047)	0.15*** (0.054)	0.13*** (0.049)
$\beta : [PostPNTR_t \times NTRGap_j] \times \ln(pop_c)$	-0.11*** (0.033)	-0.11*** (0.033)	-0.10*** (0.028)	-0.10*** (0.035)	-0.09*** (0.029)
Industry FE	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓
City FE			✓		✓
$\ln(pop_c)$	✓	✓	(✓)	✓	(✓)
$\ln(pop_c) \times PostPNTR_t$	✓	✓	✓	✓	✓
$\ln(pop_c) \times NTRGap_j$	✓	✓	✓	✓	✓
City controls		✓	(✓)	✓	(✓)
City controls $\times PostPNTR_t$				✓	✓
City controls $\times NTRGap_j$				✓	✓
City controls $\times [PostPNTR_t \times NTRGap_j]$				✓	✓
Observations	36,685	36,433	36,682	36,433	36,430
Pseudo R^2	0.57	0.58	0.65	0.58	0.66

Notes: The table presents evidence for Proposition 5, which predicts that the local trade elasticity is smaller in larger cities. This is captured by the negative coefficient β on the triple interaction. $PostPNTR_t$ is a dummy that takes on value one for $t \geq 2001$, and $NTRGap_j$ is the counterfactual increase in export tariffs for sector j from China to the US that would have occurred if China's normal trade relations (NTR) status had lapsed, rather than becoming permanent (PNTR) in 2001. Thus, a higher $NTRGap_j$ reflects a larger reduction in tariffs, relative to the counterfactual. The regression uses the standardized $NTRGap_j$, with mean zero and std one. 'City Controls' include distance to the coast, distance to the border, and capital-to-labor ratio (all in logs). Standard errors clustered at the city and at the industry level in parentheses. Key: *** significant at 1%; ** 5%; * 10%.

Table 7: Export Intensity and City Size: Decomposition in Various Models vs. Data

	(1) $\ln \left[\begin{array}{c} \text{Export} \\ \text{intensity} \end{array} \right]$	(2) $\ln \left[\begin{array}{c} \text{Intensive} \\ \text{margin} \end{array} \right]$	(3) $\ln \left[\begin{array}{c} \text{Extensive} \\ \text{margin} \end{array} \right]$
<i>Panel A: China</i>			
(1) Data	0.272*** (0.047)	-0.182*** (0.033)	0.454*** (0.045)
(2) Baseline Model [†]	0.202*** (0.027)	0.096** (0.009)	0.182*** (0.023)
(3) Homogeneous firms model [‡]	0.022** (0.010)	0.022** (0.010)	0.000 (0.000)
(4) Baseline without internal trade costs	0.137*** (0.018)	0.000 (0.000)	0.137*** (0.018)
(5) Baseline with common shape parameter	0.024** (0.011)	0.021** (0.010)	0.002** (0.001)
<i>Panel B: France</i>			
(1) Data	0.185*** (0.041)	0.018 (0.033)	0.167*** (0.026)
(2) Baseline Model [†]	0.167*** (0.022)	0.007 (0.008)	0.161*** (0.018)
(3) Homogeneous firms model [‡]	0.007 (0.009)	0.007 (0.008)	0.000 (0.000)
(4) Baseline without internal trade costs	0.150*** (0.016)	0.000 (0.000)	0.150*** (0.016)
(5) Baseline with common shape parameter	0.008 (0.009)	0.007 (0.008)	0.001 (0.001)

Notes: The table decomposes cities' overall export intensity (export sales over total sales, column 1) into two components (all in logs): Export intensity of exporters – the intensive margin (col 2) and the total sales ratio of exporting firms to all firms – the extensive margin (col 3). Cities are defined in terms of Metropolitan Areas and Commuting Zones for China and France, respectively; the analysis considers cities with positive exports and at least 250 firms. All specifications include the 'Geography Controls' listed in the note to Table 2. Robust standard errors are in parentheses. Key: *** significant at 1%; ** 5%; * 10%.

[†] City-specific Pareto shape and scale parameters.

[‡] City-specific average productivity: firms are *homogeneous* within cities, while allowing for variation in productivity levels *across* cities.

Table 8: Model Counterfactual: Trade Liberalization

	(1)	(2)	(3)	(4)
	Aggregate Welfare	Population Share 90-10 Ratio	Nominal Wage 90-10 Ratio	Real Income 90-10 Ratio
<i>Panel A: China</i>				
(1) Baseline Model [†]	4.4%	10.0%	1.9%	2.4%
(2) Homogeneous firms model [‡]	5.5%	9.3%	0.8%	1.6%
(3) Baseline without internal trade costs	4.4%	2.5%	1.4%	1.2%
(4) Baseline with common shape parameter	5.2%	9.2%	0.8%	1.6%
(5) Melitz – No internal geography	6.2%	—	—	—
<i>Panel B: France</i>				
(1) Baseline Model [†]	4.3%	6.9%	2.3%	2.2%
(2) Homogeneous firms model [‡]	4.9%	2.6%	0.7%	1.1%
(3) Baseline without internal trade costs	4.4%	4.6%	1.7%	1.2%
(4) Baseline with common shape parameter	4.6%	2.6%	0.7%	1.1%
(5) Melitz – No internal geography	5.7%	—	—	—

Notes: The table shows the estimated effects of moving from autarky to trade in terms of aggregate welfare (col 1), spatial differences in population and wages (col 2-3), and overall wage inequality (col 4) for different versions of the model presented in Section 5.5 and discussed in more detail in Appendix E. Row 1 presents results for our baseline model that allows for city-varying productivity dispersion, with city-specific Pareto shape Parameters. Row 2 presents results for a version of the model with homogeneous firms. Row 5 presents results for a model that does not have any domestic trade costs and imposes that the Pareto *shape* parameter is equal across cities, while allowing for differences in the *scale* parameter of the local productivity distributions.

[†] City-specific Pareto shape and scale parameters.

[‡] City-specific average productivity: firms are *homogeneous* within cities, while allowing for variation in productivity levels *across* cities.

Table 9: Model Counterfactual: Less Restrictive Land Use Regulation in the Baseline Model

	(1)	(2)	(3)	(4)	(5)
	Population Share 90-10 Ratio	Aggregate Export	Aggregate Sales	Export Intensity	Aggregate Productivity
(1) China	73.9%	4.8%	4.3%	0.05 p.p.	6.0%
(2) France	46.0%	0.8%	1.3%	-0.06 p.p.	1.6%

Notes: The table shows the estimated effects of increasing the housing supply elasticity from the 25th to the 75th percentile of previous estimates (see Section 5.5 for details) in terms of aggregate welfare (col 1), spatial concentration (col 2), aggregate exports (col 3), aggregate sales (col 4), export intensity (col 5), and aggregate productivity (col 5) for the baseline model.