CESifo Working Papers
www.cesifo.org/wp

Geography, Non-Homotheticity, and Industrialization: A Quantitative Analysis

Holger Breinlich
Alejandro Cuñat

CESifo Working Paper No. 3927
Category 8: Trade Policy
August 2012

CESifo *Working Paper No. 3927*

Geography, Non-Homotheticity, and Industrialization: A Quantitative Analysis

Abstract

We propose a quantitative framework for the analysis of industrialization in which specialization in manufacturing or agriculture is driven by comparative advantage and non-homothetic preferences. Countries are integrated through trade but trade is not costless and geographic position matters. We use a number of analytical examples and a multi-country calibration to explain two important empirical regularities: (i) there is a strong positive correlation between proximity to large markets and levels of manufacturing activity; (ii) there is a positive correlation between the ratio of agricultural to manufacturing productivity and shares of manufacturing in GDP. Our calibrated model replicates these facts and also provides a better fit to cross-sectional data on manufacturing shares than frameworks which ignore the role of trade costs or non-homotheticity. We use the calibrated model to quantitatively analyze the effect of increases in agricultural productivity and a further lowering of trade barriers.

JEL-Code: F110, F120, F140, O140.

Keywords: industrialization, economic geography, international trade.

Holger Breinlich
Department of Economics
University of Essex
Wivenhoe Park
UK – Colchester CO4 3SQ
hbrein@essex.ac.uk

Alejandro Cuñat
Department of Economics
University of Vienna
Hohenstaufengasse 9
Austria – 1010 Vienna
alejandro.cunat@univie.ac.at

July 2012
This paper is partly based on the unpublished 2005 paper “Economic Geography and Industrialization” which was chapter 2 of Breinlich’s PhD dissertation. We are grateful to the editor, two anonymous referees, Harald Fadinger, Gabriel Felbermayr, Jon Temple and seminar participants in Copenhagen, Mannheim, Munich and Vienna for helpful suggestions. Stephen Redding, Anthony Venables and Silvana Tenreyro provided very useful comments on the earlier PhD chapter. All remaining errors are ours. Cuñat gratefully acknowledges financial support by the Austrian Science Fund (FWF #AP23424-G11), and the hospitality of CES-ifo while revising this paper.

1 Introduction

This paper revisits a question with a long tradition in development economics. What explains industrialization, *i.e.*, the decline of agriculture's share in GDP and the corresponding rise of manufacturing (and later services)? Why do we observe such substantial differences in levels of industrialization around the world?

The literature on structural change has proposed a number of theories to explain these phenomena. The most influential approaches focus on differences in the income elasticity of demand across sectors (*e.g.*, Murphy *et al.* (1989b); Kongsamut *et al.* (2001)), sector-biased productivity growth (e.g., Ngai and Pissarides (2007)), or a combination of both (e.g., Caselli and Coleman (2001); Duarte and Restuccia (2010)). Traditionally, these approaches have analyzed closed-economy models. More recently, several authors have provided extensions to open-economy settings and have shown that additional forces, such as comparative advantage, become relevant in such models and can substantially alter the results from the closed-economy literature.[1]

While analyzing the phenomenon of industrialization in open-economy models seems natural in today's integrated world economy, it is important to realize that trade is not costless and that distance and thus the geographic position of countries still matters. Indeed, a large literature in international trade and economic geography has highlighted the links between industrial specialization, trade and geography. For example, authors such as Krugman (1980), Krugman and Helpman (1985), Davis and Weinstein (2003) or Behrens et al. (2009), to name but a few, have shown how market size and relative geographic position shape industrial specialization patterns. Eaton and Kortum (2002), Fieler (2011) and Eaton, Kortum and Kramarz (2011), among others, have used fully parameterized multi-country models with trade costs to explain observed trade patterns and to quantify the gains from trade.

In this paper, we propose a quantitative framework for the analysis of industrialization in which countries are integrated through trade, but where trade is not costless and geographic position matters. Building on a set of theoretical mechanisms well known in the trade and economic geography literature, we construct a multi-country model with costly trade augmented with a key ingredient of structural change models: non-homothetic preferences that lead to an income elasticity of demand for manufacturing higher than for agriculture. We argue that such a framework is useful to understand two important empirical regularities, and provides a better fit to cross-sectional data on manufacturing shares than frameworks which ignore the role of economic geography.

The first observation that motivates our choice of framework is that proximity to foreign sources of demand seems to matter for industrialization. For example, it has long been noted that Hong Kong, Singapore, and Taiwan not only benefitted from an outward-oriented trade policy but also close proximity to the large Japanese market. A cursory look at the data suggests that distance to foreign markets has a more general relevance: Figure 1 plots the manufacturing

[1] Important recent contributions in the literature on growth, industrialization and structural change in open economies or many-country models with free trade include Coleman (2007), Galor and Mountford (2008), Matsuyama (2009), and Yi and Zhang (2010).

share in GDP against the minimum distance to the European Union, Japan and the U.S. for a cross-section of developing countries in 2000.[2] The figure shows that developing economies close to one of these main markets of the world have higher levels of industrialization as measured by manufacturing's share in GDP.

Our second observation is that a standard proxy for comparative advantage in agriculture, labor productivity in agriculture relative to manufacturing, is either positively or not at all correlated with manufacturing shares in the developing world. Figure 2 plots these two variables against each other for a cross-section of developing countries for the year 2000.[3] The fitted line has a positive, albeit statistically insignificant slope. As we show in our more detailed econometric analysis in Section 2, extending the sample to include more countries and years leaves this positive correlation intact and actually makes it statistically significant as well.

By construction, the first of these two facts cannot be explained by closed-economy models of industrialization. But it also sits uneasily with open-economy models with free trade, in which geographic position is irrelevant. Figure 2 is even more puzzling for open-economy theories of industrialization that stress the importance of comparative advantage. If countries are indeed integrated through trade and comparative advantage forces are active, should we not expect to find a *negative* correlation in the data?

In fact, both observations arise naturally in a model nesting non-homothetic preferences in a multi-country comparative-advantage trade model with positive but finite trade costs. In the model we propose below, developing countries closer to foreign sources of demand will experience higher demand for both the agricultural and manufacturing goods they produce than more distant countries, ceteris paribus. Following contributions to the international trade and economic geography literature (e.g., Hanson and Xiang (2004) or Murata (2008)), we outline conditions under which this translates into higher manufacturing shares in GDP. Most importantly, higher overall demand will lead to higher wages which, in the presence of non-homothetic preferences combined with positive trade costs, will shift local production towards the manufacturing sector. This effect is further reinforced if manufactured products are more differentiated than agricultural products. Trade costs for agricultural products also hamper the comparative-advantage mechanism put forward by free-trade models. High agricultural productivity leads to higher wages which, again because of the combination of agricultural trade costs and non-homothetic preferences, leads countries to specialize in manufacturing (we call this the "relative-demand effect" of agricultural productivity). The standard comparative-advantage effect, which would drive specialization patterns in the opposite direction, is also present but can be overcompensated by the relative-demand effect for intermediate levels of trade costs.

[2] We use the Netherlands as the approximate geographic centre of the European Union in Figure 1. Developing countries are defined as countries belonging to the income categories "low", "lower middle" and "upper middle" published by the World Bank (corresponding to less than 9,265 USD in 1999). The simple OLS regression underlying the fitted line in Figure 1 yields a negative slope coefficient which is statistically significant at the 1% level.

[3] Developing countries are defined as in footnote 2. Labor productivity is measured as value added per worker in agriculture and manufacturing, respectively, where value added is corrected for cross-country price differences using sector-specific PPP exchange rates (also see Section 2 and Appendix B). This is the proxy of choice in many studies of Ricardian comparative advantage, *e.g.* Golub and Hsieh (2000).

Having shown that our model can, in principle, explain our stylized facts, we proceed to a calibration of our model based on data from 107 developed and developing countries for the year 2000. The purpose of this calibration is threefold.

First, we show that the model also matches our stylized facts for empirically plausible parameter values. We choose parameters to match international trade and expenditure data, and demonstrate that our calibrated model generates the same positive correlations observed in the data between access to foreign markets and comparative advantage in agriculture, on the one hand, and manufacturing shares, on the other hand. Crucially, this is not true when we constrain our trade cost estimates to be equal to zero ('free trade') or infinitely high ('autarky'), or when we change parameter values to eliminate the non-homotheticity from the preferences in our model.

Secondly, we show that allowing for non-homotheticity and positive but finite levels of trade costs also improves the model's predictive power (in terms of matching observed and predicted GDP shares of manufacturing relative to agriculture) as opposed to autarky and free trade, or compared to models without non-homotheticity in demand.

Finally, having demonstrated the empirical relevance of our calibrated model, we use it to perform a number of counterfactual experiments. In a first set of experiments, we measure how the impact of increases in agricultural productivity on manufacturing activity depends on a country's access to foreign markets. Interestingly, we find that the potential deindustrializing comparative advantage effect of such increases is overcompensated by the relative demand effect highlighted above for almost all countries in our calibrated model. In a second set of experiments, we compare the impact of global and regional reductions in trade barriers on levels of industrialization. While global reductions lead to stronger increases in real income and thus (via non-homothetic preferences) to higher levels of manufacturing activity, they also drive countries to specialize according to their comparative advantage. This lowers manufacturing shares in countries with relatively high agricultural productivity. On the other hand, regional reductions in trade barriers do not raise income levels by as much, but the comparative advantage effect of such reductions is also more muted. This is because, empirically, differences in comparative advantage tend to be smaller across neighbouring countries.

Our paper is related to two sets of contributions in the literature. The first is the literature on industrialization and structural change already discussed above. Within this literature, our paper relates most closely to approaches relying on differences in the income elasticity of demand across sectors for explaining structural change (e.g., Murphy *et al.* (1989b); Kongsamut *et al.* (2001)). We note, however, that our stylized facts also cannot be easily accounted for in existing theoretical frameworks based on sector-biased productivity growth such as the ones cited above.

We also draw on a large literature in international trade and economic geography concerned with the effects of comparative advantage and relative location on specialization patterns in multi-country frameworks with trade costs. For example, the first of our stylized facts directly echoes the empirical finding from the literature on home-market effects that the geographical distribution of demand matters for local production patterns (e.g., Davis and Weinstein, 2003). The explanation we provide for our empirical findings is also based on a number of well-known theoretical mechanisms. For example, Behrens *et al.* (2009) analyze home-market effects in a multiple-location setup with trade costs. Hanson and Xiang (2004) show that industries with more differentiated products tend to be more concentrated in large countries than industries with less differentiated products. Epifani and Gancia (2006) derive a similar result in models with increasing returns to scale, imperfect competition, and variable markups, in which the larger the size of the market, the higher the specialization in industries with high product differentiation.

There are also many trade models with non-homotheticities on the demand side.[4] Simplifying this literature in a rather crude way, with non-homothetic preferences international differences in income and income distribution affect relative demand patterns; these have an effect on trade patterns by affecting relative excess demands (for given production structures) and by biasing the production structures of countries towards their high-demand industries (in the presence of trade costs). In Markusen (1986), for example, rich trading partners consume and specialize in industries subject to increasing returns (that is, manufacturing) and as a result more North-North (intra-industry) trade takes place with higher incomes or with larger degrees of non-homotheticity in preferences. Mitra and Trindade (2005) show how non-homothetic preferences combined with differences in income distribution can revert patterns of comparative advantage in classic trade models; in general, however, consumption and specialization patterns run in parallel in the data, which suggests the need to introduce trade costs in demand-driven trade models. In Fajgelbaum, Grossman and Helpman (2011), a model with trade costs, richer countries specialize in goods with higher income demand elasticity due to the effect of higher income on relative demands. Fieler (2011) combines non-homothetic preferences and Ricardian comparative advantage in a many-country model with trade costs, thus generalizing the Eaton and Kortum (2002) framework.[5] Finally, Murata (2008) is a contribution from the New Economic Geography literature that uses non-homothetic preferences in a core-periphery model with trade costs based on Krugman (1991).

Our contribution relative to the above sets of literature is twofold. First, we show how insights from the international trade and economic geography literature can be applied to the modelling of cross-sectional patterns in levels of industrialization. While this is a relatively straightforward extension of existing results, it nicely complements existing work in the structural change literature, which is mostly based on dynamic frameworks but ignores geographic aspects of industrialization.[6] More specifically, our results show that home-market effects are also a powerful factor in explaining the cross-country variation in manufacturing shares, in addition to their impact on

[4]Many of these references trace back their ideas to Linder (1961). As we explain in more detail below, most of the literature consists of highly stylized models that cannot be easily calibrated in a multi-country setting (Fieler (2011) is an exception, again more on this below).

[5]See Matsuyama (2000) for a "traditional" Ricardian model with a continuum of goods and non-homothetic preferences.

[6]Puga and Venables (1999) and Murata (2008) also use theoretical mechanisms from the new economic geography literature to explain aspects of structural change. However, their focus is, respectively, on the sequential spread of industries across countries, and on jointly accounting for shifts in expenditure and labor allocation from agricultural to non-agricultural goods and the emergence of urban agglomeration. Both papers also use more stylized setups and do not directly connect their models to the data as we do. Desmet and Rossi-Hansberg (2011) present a spatial endogenous growth model which they use to quantitatively explain structural change in the United States. But they rely on sector-biased productivity growth instead of non-homothetic preferences, and focus on the transition from manufacturing to services (which is less relevant for developing countries).

finer specialization patterns as documented in, for example, Davis and Weinstein (2003). Indeed, as we explain below, the explanatory power of our centrality measures for the cross-country variation in manufacturing shares is of comparable magnitude to per-capita income, one of the key explanatory variables in the empirical literature on structural change (e.g., Syrquin and Chenery, 1989). Likewise, our second stylized fact has important implications for the modelling of comparative advantage forces in open-economy approaches to structural change. While a positive or insignificant correlation between comparative advantage in agriculture and manufacturing shares arises naturally from the theoretical mechanisms discussed above, we are not aware of any previous research that documents this correlation empirically for a broad cross-section of developing countries.

Secondly, our paper seems to be the first attempt to quantify the importance of trade costs and non-homotheticity in shaping broad sectoral specialization patterns in multi-country settings with trade costs, which is relevant for both sets of literature discussed above. In particular, we show that the effects of trade costs and non-homotheticity highlighted in the existing literature are quantitatively important and of a magnitude sufficient to explain some of the key patterns in the data. Put differently, we show that the theoretical mechanisms outlined above are not only relevant in principle, but also under plausibly calibrated parameter values. Building a fully quantified model also allows us to carry out some interesting counterfactual simulations for which the magnitude of parameter values is important and which is an additional contribution of our paper. Among recent quantitative work in international trade and economic geography, Fieler (2011) is the only contribution we are aware of which also introduces non-homothetic preferences into a multi-country setup with trade costs. However, her focus is different from ours in that she is trying to explain differences in trade patterns and volume across rich and poor countries, and there are also some important differences regarding the theoretical framework and the estimation strategy used in her and our paper.[7]

The remainder of this paper is organized as follows. Section 2 shows that the two correlations displayed in Figures 1 and 2 are robust to changes in sample composition and to the inclusion of proxies for local demand and other domestic factors. Section 3 develops a multi-country model with trade costs. Using a number of analytical examples, we use it to provide an explanation for the correlations highlighted in this introduction. In Section 4, we calibrate the model to match international trade and expenditure data and show that this calibrated version generates the same correlations as in Figures 1 and 2 and Section 2. We also conduct a number of counterfactual experiments, which further illustrate the importance of trade costs and non-homotheticity in shaping cross-sectional patterns of industrialization in the developing world. Finally, Section 5 concludes.

[7] For example, she relies on a probabilistic representation of technologies as in Eaton and Kortum (2002) instead of Armington-based product differentiation as in our paper (i.e., heterogeneity within goods is introduced from the supply rather than from the demand side). Her estimation strategy is also fundamentally different from ours in that all parameters apart from country-specific technology parameters are chosen to match bilateral trade flows.

2 Empirical Evidence

In this section, we examine the robustness of the correlations from the introduction through variations in sample composition and by including a number of control variables. The correlations presented in the following are the ones we will aim at reproducing in our calibration exercise. Our full econometric specification will be

$$ltShareM_{lt} = \alpha + d_t + \beta_1 RP_{lt} + \beta_2 CEN_{lt} + \beta_3 AP_{lt} + \beta_4 POP_{lt} + \varepsilon_{lt}, \tag{1}$$

where RP_{lt} is relative productivity (of agriculture to manufacturing) and CEN_{lt} the 'centrality' of country l, *i.e.*, its access to foreign markets (to be defined below). AP_{lt} denotes agricultural productivity, POP_{lt} the population size of country l, and d_t is a full set of year fixed effects. The dependent variable is the logistic transformation of a country's share of manufacturing value added in GDP. We use a logistic transformation to account for the fact the manufacturing share is limited to a range between 0 and 1.[8] Concerning the regressors, we discuss the choice of suitable empirical proxies in turn. Additional details on the data and their sources as well as a list of countries used in the regressions below are contained in Appendix A.

Keeping in line with existing studies on Ricardian comparative advantage (*e.g.,* Golub and Hsieh, 2000), we use labor productivity as a proxy for productivity. In contrast to total factor productivity, this has the advantage of considerably increasing the number of available observations. We measure labor productivity as value added per worker in agriculture and manufacturing, respectively. Importantly, we use data on sector-specific purchasing power parities to strip out the cross-country variation in prices from the relative productivity data, so that the remaining variation more closely reflects physical productivity differences (see Appendix B for additional details). This ensures the comparability of our empirical exercises (both here and in the quantitative examples below) with our theoretical framework.

We measure country l's centrality (CEN_{lt}) as the sum of all other countries' GNP, weighted by the inverse of bilateral distances, which are taken to proxy for trade costs between locations:

$$CEN_l = \sum_{j \neq l} GNP_j \times dist_{jl}^{-1}. \tag{2}$$

This specification reflects the basic intuition of our discussion. What matters is centrality in an economic geography sense, that is proximity to markets for domestic products. Of course, the above centrality index is closely related to the concept of market potential first proposed by Harris (1954), which has been frequently used in both geography and – more recently – in economics. A number of studies have demonstrated that this simple proxy has strong explanatory power and yields results very similar to more complex approaches that estimate trade costs from trade flow

[8] Using untransformed manufacturing shares instead does not change any of the qualitative results reported below. We have also experimented with including the share of services in GDP as an additional control variable, again without finding any significant changes in the other coefficient estimates (both sets of results are available from the authors upon request).

gravity equations (see, for example, Head and Mayer (2006), or Breinlich (2006)).[9]

As additional control variables, we also include agricultural productiviy (AP) to account for the pro-industrializing relative-demand effect discussed above, and population size (POP) as an additional proxy for the extent of the domestic market. We have data for all the required variables for 112 countries in 2000. Keeping in line with the focus of this paper on the industrialization of developing countries, however, we exclude high-income countries from our regression sample (although of course all available countries are used to calculate the centrality measure).[10] In our robustness checks, we will also briefly present results for the full sample.[11]

In Table 1, we present a number of univariate correlations between the logistic transformation of manufacturing shares and our proxies for comparative advantage (relative productivity, RP) and centrality. Columns 1-2 replicate the correlations from Figures 1 and 2 and show that using a logistic transformation of manufacturing shares as the dependent variable leads to similar results. In column 3, we use our more sophisticated measure of centrality (2). Note that we would now expect to find a positive and significant sign, which is indeed what we do. We also note that both measures of centrality seem to be important determinants of levels of industrialization. They explain around 10% of the cross-sectional variation of manufacturing shares in our sample. This is comparable in magnitude to *per-capita* income, whose positive correlation with manufacturing shares in the initial phase of development is a key variable in much of the existing empirical literature on cross-country patterns of industrialization (e.g., Syrquin and Chenery (1989)).

In columns 4-8, we undertake a first series of robustness checks. One concern with our baseline centrality measure is that a large proportion (approximately 40%) of bilateral trade flows between the countries used in the calculation of CEN_l are zero. This raises the question of whether we are simply adding up the GNPs of unrelated foreign countries in (2). To address this concern, we also present results for a centrality measure which only includes the GNP of a foreign country j in the calculation of our centrality measure for country l if we observe non-zero trade flows between j and l. Column 4 shows that the results for this alternative measure are basically identical to our baseline results. Indeed, the correlation between the two centrality measures in logs is 99%.[12] This seems to be due to the fact that zero trade flows are less common in the data if the trading partner is close by and/or has a high GNP. But these trading partners are also the countries which make the largest contribution to the calculation of our original centrality measure. Put differently, the countries we are dropping now only made small contributions to the original measure and dropping them does not significantly change its value nor the ranking of countries along the centrality measure.

Column 5-6 include a dummy for China and the South-East Asian economies of Korea, Thai-

[9] Using a nonstructural measure also seems to be better in line with the more explorative character of this section.

[10] We use the World Bank's income classification and exclude all countries with gross national income per capita in excess of 9,265 USD in 1999 ("high income countries").

[11] See footnote 32 in Section 4.2 and Appendix Table A.2.

[12] Note that we had to drop three countries for the new measure (Botwana, Lesotho and Namibia) for which we do not have data on bilateral trade flows. Running our original regression from column 3 on the smaller sample of 80 countries again yields very similar results (the reported 99% correlation is for this sample).

land, Malaysia, Indonesia and the Philippines. These countries are arguably special cases because of their very successful export-oriented industrialization strategies and are also potentially influential outliers in both Figures 1 and 2. The corresponding dummy variable (not reported) is indeed positive and highly significant but the coefficient on our centrality measure remains almost unchanged. The positive correlation between manufacturing shares and relative productivity is increased and becomes statistically significant. In columns 7 and 8, we present results for additional years for which comparable cross-sectional data on relative productivities is available (1980 and 2000, yielding a unbalanced panel of 256 observations in total). Again, using these additional data makes the results from columns 1 and 3 stronger.[13]

In Table 2, we gradually build up our results to the full specification (1). In column 1 we include population size; column 2 uses agricultural productivity as an additional regressor; and column 3 includes both population and agricultural productivity.[14] In column 4, we drop agricultural productivity and replace it with *per-capita* GDP. *Per-capita* GDP helps controlling for the purchasing power of the local population, skill levels, and other potentially confounding factors. Note, however, that it is very highly correlated with agricultural productivity so that in practice both variables are likely to pick up the influence of similar omitted variables. The high correlation also makes the inclusion of both variables in the same regression impossible.[15] In columns 5-8, we again use our larger sample for the years 1980, 1990 and 2000.

Three main insights arise from these regressions. First, proxies for the size of the domestic market are strongly positively correlated with levels of industrialization, as was to be expected from prior results in the literature. Second, centrality retains its positive and significant influence throughout. Third, comparative advantage in agriculture has a positive and significant effect on industrialization whenever we do not control for absolute agricultural productivity, and an insignificant effect whenever we do. This suggests that relative productivity might be picking up the influence of absolute productivity levels in agriculture.

Limited data availability for relative and absolute agricultural productivity prevents us from estimating specification (1) for a yet larger sample. In columns 9-11, we exclude these variables which increases the sample size more than tenfold since we can now use observations for every year from 1980 to 2005. This allows us to provide some further results on the importance of centrality for industrialization by running variations of the following specification:

$$ltShareM_{lt} = \alpha + d_t + d_l + \delta_1 CEN_{lt} + \delta_2 PCGDP_{lt} + \delta_3 POP_{lt} + \varepsilon_{lt}, \tag{3}$$

where $PCGDP_{lt}$ denotes *per-capita* GDP and d_t and d_l are a full set of time and country fixed

[13] Note that our PPP data do not have sufficient country coverage for these earlier years to correct our productivity data for price differences. For example, there are only around 60 countries in the 1985 and 1980 waves of the ICP and, moreover, developing countries are underrepresented in these years. Thus, we use market exchange rates to convert the value added data to a common currency (USD with base year 2000). However, results for the year 2000, for which we can do both PPP and market exchange rate conversions, are qualitatively similar across both approaches.

[14] As before, we measure agricultural productivity as value added per worker in agriculture, adjusted for cross-country price variation (see above and Appendix B for details).

[15] The correlations of the variables in logs is 84% in our sample.

effects. Column 9 of Table 2 reports results for an OLS regression pooled over the period 1980-2005 with year dummies only. Column 10 estimates the full specification (3) by including country fixed effects, thus eliminating any time-invariant heterogeneity across countries from our correlations. Column 11 uses long first differences between 1980 and 2005. All regressions give a similar picture as the results for the smaller sample: both the size of the domestic market and access to foreign markets are positively correlated with levels of industrialization. If anything, controlling for country-specific effects in columns 10 and 11 implies an even stronger role for centrality.[16]

3 The Model

We now outline a simple multi-country model in which the cross-country variation in manufacturing and agricultural GDP shares is driven by non-homothetic preferences, and absolute and relative sectoral productivities. Because we will allow for positive but finite trade costs, this model will be able to generate the correlations just presented.

3.1 Model Setup and Equilibrium

Consider a world with countries $j = 1, ..., R$, each with L_j consumers, each of which supplies one unit of labor inelastically. There are two sectors, agriculture and manufacturing; we assume perfect labor mobility between sectors, and no international labor mobility. As we discuss in Section 4.2 and Appendix D below, adding a third, non-tradable sector (*i.e.*, services) complicates the analysis but yields similar results, both qualitatively and (in our calibration) quantitatively. Thus, for the sake of simplicity we abstract from the services sector for most of our analysis.

Preferences are identical across countries. Country-j individuals maximize a Stone-Geary utility function over consumption of an agricultural and a manufacturing composite good:

$$U_j = \alpha \ln(M_j) + (1-\alpha)\ln(A_j - \underline{A}), \tag{4}$$

$\alpha \in (0,1)$, where

$$M_j = \left[\sum_{l=1}^{R} m_{lj}^{(\sigma_M-1)/\sigma_M}\right]^{\sigma_M/(\sigma_M-1)}, \tag{5}$$

$$A_j = \left[\sum_{l=1}^{R} a_{lj}^{(\sigma_A-1)/\sigma_A}\right]^{\sigma_A/(\sigma_A-1)}. \tag{6}$$

Both M_j and A_j are Armington aggregators of country-specific varieties: every country is assumed to produce one differentiated variety.[17] M_j is consumption of the manufacturing composite and

[16] Omitting relative and absolute agricultural productivity does of course lead to an omitted variable bias. To verify the likely magnitude of this bias, we estimated both (1) and (3) on the same samples used in columns 1-4 and 5-8. Comparing the coefficient on CEN in these regressions does indeed suggest that omitting AP and RP leads to an upward bias, albeit a small one.

[17] The Armington assumption ensures that all countries consume all varieties provided trade costs and elasticities

m_{lj} is the amount of the variety produced in l that is consumed by an individual consumer in j. Similarly, A_j is consumption of the agricultural composite and a_{lj} is the amount of the variety produced in l that is consumed by an individual consumer in j. The elasticities of substitution between varieties are constant at $\sigma_M, \sigma_A > 1$.

$\underline{A} > 0$ denotes minimal consumption of agricultural goods, *i.e.*, the subsistence level. These preferences guarantee that above the level $\underline{A}$ the expenditure share of agricultural goods declines with rising *per-capita* income. This is the so-called Engel's law, which has strong empirical foundations (see, for example, Crafts (1980)). We assume $\underline{A} < \theta_{Al}$ for all l, where θ_{Al} is agricultural productivity in country l (to be defined below). This assumption guarantees that *per-capita* income in each country is sufficient to reach the subsistence level. Thus, at least some expenditure will be devoted to manufacturing products.

Below we impose enough structure so that labor is the only source of income. The individual's budget constraint in country j is therefore given by $P_{Mj}M_j + P_{Aj}A_j = w_j$, with w_j denoting the wage in j, equal across sectors. P_{Mj} and P_{Aj} are price indices for the manufacturing and agricultural composite goods. Prices paid for the different products in the importing location j, p_{Mlj} and p_{Alj}, consist of the mill price charged in country l plus industry-specific bilateral trade costs $T^M_{lj}, T^A_{lj} \geq 1$. ($T^M_{jj} = T^A_{jj} = 1$ for all j.) These trade costs are of the iceberg-type form: for every unit of a good that is shipped from l to j, only $1/T_{lj}$ arrive while the rest "melts" *en route*.

By assumption, each country produces a differentiated variety of the manufacturing and the agricultural good. We also assume that firms operate under constant returns to scale, and use labor as the only input. Since goods are just differentiated by origin and there are constant returns to scale, we can safely assume many firms in each sector-country so that industries are perfectly competitive. The amount of labor employed in manufacturing in country l is denoted by L_{Ml}, and supply of the local variety is $m_l = \theta_{Ml}L_{Ml}$, where θ_{Ml} denotes productivity in manufacturing in country l. The amount of labor employed in agriculture in country l is denoted by L_{Al}, and supply of the local variety is $a_l = \theta_{Al}L_{Al}$. Productivity levels are allowed to vary across countries and sectors. Positive production implies f.o.b. prices equal the cost of producing one unit of output: $p_{Ml} = w_l/\theta_{Ml}$ and $p_{Al} = w_l/\theta_{Al}$.

Equilibrium in the manufacturing and agricultural goods markets requires that demand for each Armington variety equals its supply. With goods markets clearing, we can express labor demand as a function of the vector of wages of all countries. Full employment requires $L_{Ml}+L_{Al} = L_l$., which can be rewritten as

of substitution are finite. This implies that all countries have diversified production structures in equilibrium, which is the empirically relevant case and makes the calibration of the model significantly easier. It has the drawback, however, that it rules out zero trade flows between country pairs, which is not true in the data. Trade theory addresses this issue by assuming demand systems that yield finite reservation prices, such as Melitz and Ottaviano (2008) and Behrens, Mion, Murata and Südekum (2009). Calibrating such models for our sample of countries represents an important challenge due to data availability issues. Since bilateral trade flows are not our main object of study in this paper, we leave this problem for future research.

$$\theta_{Ml}^{\sigma_M-1} w_l^{-\sigma_M} \left[\sum_{j=1}^{R} \left(T_{lj}^{M}\right)^{1-\sigma_M} P_{Mj}^{\sigma_M-1} E_{Mj} \right] + \theta_{Al}^{\sigma_A-1} w_l^{-\sigma_A} \left[\sum_{j=1}^{R} \left(T_{lj}^{A}\right)^{1-\sigma_A} P_{Aj}^{\sigma_A-1} E_{Aj} \right] = L_l. \quad (7)$$

These are R non-linear equations in the R wage rates, determining the vector of wages and subsequently all other equilibrium variables of the model. For the analysis below, we also define the share of manufacturing in GDP, which in the following is also referred to as the level of industrialization of a location:

$$ShareM_l = \frac{p_{Ml} m_l}{p_{Ml} m_l + p_{Al} a_l} = \frac{w_l L_{Ml}}{w_l L_{Ml} + w_l L_{Al}} = \frac{L_{Ml}}{L_l}. \quad \text{(MS)}$$

3.2 Analysis

Before proceeding to the quantification of the model just outlined, we briefly discuss a number of analytical results which help to build intuition for the full-scale multi-country calibration which follows in the next section. We start by discussing the two polar cases of infinite and zero trade costs ("closed economy" and "free trade"), and then move on to the general case with positive trade costs.

3.2.1 Closed Economy

With infinitely high levels of trade costs, *i.e.*, under autarky, it is easy to show that the expression for the share of manufacturing in GDP simplifies to

$$ShareM_l = \alpha \left(1 - \frac{\underline{A}}{\theta_{Al}} \right). \quad (MS_{AUT})$$

As is apparent from equation (MS_{AUT}), the manufacturing share in GDP increases with agricultural productivity. Non-homothetic preferences (due to the positive subsistence consumption level in agriculture, $\underline{A} > 0$) are crucial for this result. Intuitively, the increases in *per-capita* income resulting from higher values of θ_{Al} lead to a decline in the share of subsistence consumption in total expenditure. As every unit of income above the subsistence level is spent in fixed proportions on agricultural and manufacturing varieties, the expenditure share of the latter rises. In a closed economy, this leads in turn to a shift of labor into manufacturing and an increase in $ShareM_l$. As discussed, we refer to this positive impact of agricultural productivity on industrialization as the "relative-demand effect" of agricultural productivity shocks. Very similar effects are obtained in the existing literature on structural change in closed economies (*e.g.*, Matsuyama (1992) or Murphy *et al.* (1989b)). Note that the autarky assumption renders any cross-country differences in centrality or comparative advantage completely irrelevant.

3.2.2 Free Trade

Under free trade, Ricardian comparative advantage emerges as the key factor for the determination of the level of industrialization. With costless trade, and assuming $\sigma_M = \sigma_A = \sigma,$ it is easy to show that

$$\frac{L_{Ml}/L_{Al}}{L_{Ml'}/L_{Al'}} = \left(\frac{\theta_{Ml}/\theta_{Al}}{\theta_{Ml'}/\theta_{Al'}}\right)^{\sigma-1}. \tag{8}$$

Since

$$ShareM_l = \frac{L_{Ml}}{L_l} = \frac{1}{1+(L_{Ml}/L_{Al})^{-1}}, \tag{9}$$

lower ratios θ_{Ml}/θ_{Al} imply a stronger bias towards agriculture in the production structures of countries. As in standard Ricardian models of international trade, being relatively productive in agriculture biases a country's production structure towards agriculture, thus reducing the share of manufacturing in GDP as the location specializes accordingly. We will refer to this as the "comparative advantage effect". Notice that free trade eliminates any independent influence of the productivity level θ_{Al} on industrialization via the non-homothetic preferences channel we discussed above.

3.2.3 Costly Trade

With positive but finite trade costs, a number of new effects arise in our model. We briefly discuss the underlying intuition here, and refer the reader to Appendix C for a number of analytical examples that illustrate these effects for simplified versions of the full model we will use in the calibration.

The first set of results we outline will help us understand why developing countries that are closer to major markets will have a higher manufacturing share in GDP. It is a long-standing theoretical result in international trade theory that the size of the home market matters for industrial structure (Krugman (1980), Krugman and Helpman (1985)). In a world with positive trade costs, central locations have effectively a larger market size as they are closer to sources of demand, *ceteris paribus*. Note that it does not follow immediately that such countries will specialize in manufacturing because more central locations will also experience higher demand for both agricultural and manufacturing goods relative to less central locations. Several mechanisms have been proposed in the literature for how specialization will nevertheless obtain. Two are of particular relevance here.

First, being more central raises demand for both agricultural and manufacturing goods and raises wages.[18] With non-homothetic preferences, this leads to an expansion of domestic manufacturing expenditure. If there are positive trade costs, this will translate into a domestic manufacturing share higher than in other countries, as the resulting increase in manufacturing expenditure will have a stronger effect on the domestic manufacturing good than on those produced by other countries (see Murata (2008) for a similar mechanism). Example 1 in Appendix C

[18] See Redding and Venables (2004) for empirical evidence on the positive effect of centrality on income levels.

illustrates this mechanism for a simplified version of our model with three countries (the minimum necessary to explain the role of centrality).

A second reason why centrality favors industrialization is based on different elasticities of substitution of manufacturing and agricultural products (see Hanson and Xiang, 2004). Higher wages due to a more central position lead to higher prices of both types of goods. If agricultural goods are more homogeneous than manufacturing goods (this would correspond to $\sigma_A > \sigma_M$ in our model), central locations will specialize in manufacturing, *ceteris paribus*. This is since demand for locally produced manufacturing varieties will be less sensitive to higher prices than demand for the country's agricultural variety. Example 2 in Appendix C illustrates this mechanism, again using a simplified version of our full model.

Trade costs can also affect the response of specialization patterns to changes in productivity (also compare Murata (2008)). An increase in agricultural productivity, for example, will generate a "relative-demand effect" in favor of the manufacturing industry through the non-homothetic preferences, and a "comparative-advantage effect" in favor of agriculture. Which effect dominates depends on the link between domestic expenditure and production and thus the level of trade costs. Under autarky, where consumption and production are perfectly linked, the relative-demand effect dominates, as we already saw above. Under free trade, where consumption and production are separate choices, the comparative advantage effect dominates. With intermediate values for trade costs, which effect dominates depends on parameter values (see example 3 in Appendix C).

4 Quantitative Analysis

In this section, we calibrate a multi-country version of our model using data on 107 developed and developing countries. We choose parameter values to match data on international trade flows, expenditure, productivity levels and population size (Section 4.1). We use this calibrated model to generate data on manufacturing shares and the independent variables used in the regressions in Tables 1 and 2, and show that the generated data display the same correlations between centrality, comparative advantage and manufacturing shares found in the actual data (Section 4.2). In Section 4.3, we compare our baseline calibration to three alternative parameterizations of our model – one with infinite trade costs ('autarky'), one with zero trade costs ('free trade') and one without homothetic preferences. We show that all three of these alternatives fail to reproduce our stylized facts and also perform worse in matching actual data on manufacturing shares. In a final step (Section 4.4), we use the original parameterization of our the model to shed light on two ongoing debates in the literature – the impact of higher agricultural productivity on manufacturing activity and the effects of a global vs. regional reduction in trade barriers.

4.1 Parameterization

For a calibrated version of our model, we need data on the size of countries' workforces (L_l) and productivity levels (θ_{Al}, θ_{Ml}), and values for the parameters governing substitution elasticities

(σ_A, σ_M), trade costs (T_{lj}^A, T_{lj}^M), the manufacturing expenditure share (α), and subsistence consumption ($\underline{A}$). Table 3 provides parameter estimates and a brief description of the calibration procedure and data sources used. In the following, we describe the calibration in more detail. Data requirements limit the sample to 107 countries for the year 2000, 79 of which are classified as developing and will be used in our regression analysis of the simulated data.[19]

We follow Feenstra (1994) in using variation in import quantities and prices to identify elasticities of substitution among manufacturing and agricultural varieties (σ_A, σ_M). This approach, as extended by Broda and Weinstein (2006) and Broda *et al.* (2006), has become the dominant method for estimating substitution elasticities in the international trade literature in recent years. In our setting, it has the additional advantage of building on a very similar demand structure as our paper (CES and Armington varieties), while allowing for more general supply side features. We adapt this approach to our setting by using data that correspond to our calibration exercise in terms of country coverage, time period and the definition of sectors for which we estimate elasticities. We focus on a discussion of our estimates in the following and refer the reader to Appendix E for a more detailed description of the Feenstra-Broda-Weinstein methodology and how we adapt it to our setting.

For our baseline elasticity estimates, we use cross-country trade data for the year 2000, but restrict the estimation sample to the 102 countries in our calibration sample, for which we have the necessary information on import prices and quantities.[20] We obtain $\sigma_M = 2.3$ and $\sigma_A = 2.3$. For comparison, Broda *et al.* (2006) estimate elasticities of substitution between varieties of goods produced in each of approximately 200 sectors, separately for 73 countries (rather than imposing a common elasticity as we do in accordance with our model). The median across these estimates for the 60 countries also present in our data is 3.4. Given the much higher degree of aggregation in our data (two instead of 200 sectors), our lower estimates seem plausible. This is because both economic theory and the empirical results of Broda and Weinstein (2006) and Broda *et al.* (2006) suggest that estimated elasticities should fall as the level of aggregation increases and varieties become less similar.[21]

As a robustness check, we also obtain estimates using data on imports by the U.S. from the countries in our calibration sample.[22] These data are likely to be of higher quality than the cross-country data used before (see Feenstra, Romalis and Schott, 2002), although of course we only have one importer now instead of 102. Using these data yields comparable coefficient magnitudes as before although agricultural varities are now estimated to be slightly more substitutable across

[19] See Appendix A for a list of countries included in the calibration sample. All 107 countries will be used to generate our synthetic data set as developed countries do of course play a major role in determining manufacturing shares and centrality of developing countries.

[20] Three groups of countries only report one common set of trade data, explaining the five missing observations: Botswana, Lesotho, Namibia and South Africa; Belgium and Luxembourg; and St. Lucia and St. Vincent and the Grenadines.

[21] Closer to our level of aggregation but obtained via a different methodology is the estimate by Eaton *et al.* (2011) who use French firm-level data to estimate an elasticity of substitution between individual manufacturing varieties of $\sigma_M = 1.7$.

[22] Again, we lose five countries due to aggregation in the trade data (see footnote 21), leaving us with 101 exporters (the U.S. is of course excluded as an exporter).

countries ($\sigma_M = 2.0$ and $\sigma_A = 2.6$).

Since labor is the only factor of production in our model, we proxy θ_{Ml} and θ_{Al} by labor productivity in manufacturing and agriculture, respectively. However, as already discussed in Section 2, the variation in measured labor productivity across countries and sectors that we observe in the data is driven by both differences in technological efficiency and differences in prices. That is, $lp_{.l} = VA_{.l}/L_{.l} = p_{.l}x_{.l}/L_{.l}$ in terms of our model because we abstract from intermediate inputs. Since we are only interested in $\theta_{.l} = x_{.l}/L_{.l}$, we use data on purchasing power parities for agriculture and manufacturing goods consumption from the International Comparison Program (ICP) to construct proxies for $p_{.l}$ and strip out price variation from the data (see Appendix B for details).[23]

Estimates of the trade cost matrices can be obtained via gravity equation regressions using cross-country manufacturing and agricultural trade data. To see this, note that exports in the model are:

$$\begin{aligned} X_{lj}^M &= p_{Ml}^{1-\sigma_M} \left(T_{lj}^M\right)^{1-\sigma_M} P_{Mj}^{\sigma_M-1} E_{Mj}, \\ X_{lj}^A &= p_{Al}^{1-\sigma_A} \left(T_{lj}^A\right)^{1-\sigma_A} P_{Aj}^{\sigma_A-1} E_{Aj}. \end{aligned}$$

The only bilateral variable on the right-hand side in the above expressions is trade cost T_{lj}. We proxy for these costs by $T_{lj} = dist_{lj}^{\delta_1} e^{\delta_2 d_{int,lj}}$, where $dist_{lj}$ denotes the bilateral distance between countries l and j, and δ_1 denotes the elasticity of trade cost with respect to distance. The dummy variable $d_{int,lj}$ indicates if a trade flow crosses national borders (*i.e.*, $d_{int,lj} = 1$ if $l \neq j$ and 0 if $l = j$). This is a parameterisation of trade cost which is common in the international trade literature (*e.g.*, Wei (1996)). Proxying all other variables by importer and exporter fixed effects and adding an error term, we can rewrite bilateral exports as

$$\begin{aligned} X_{lj}^M &= d_{\exp,M} \times d_{imp,M} \times dist_{lj}^{(1-\sigma_M)\delta_{M1}} e^{(1-\sigma_M)\delta_{M2} d_{int,M}} \times \varepsilon_{lj,M}, \\ X_{lj}^A &= d_{\exp,A} \times d_{imp,A} \times dist_{lj}^{(1-\sigma_A)\delta_{A1}} e^{(1-\sigma_A)\delta_{A2} d_{int,A}} \times \varepsilon_{lj,A}. \end{aligned} \tag{10}$$

We estimate (10) in its original multiplicative form via Poisson QMLE, using data from the sources listed in Table 3 and following Wei (1996) in proxying internal trade flows as domestic production (gross output) minus exports.[24] As has been pointed out by Santos-Silva and Tenreyro (2006),

[23] Echevarria (1997) uses a similar approach based on U.S. data. Note that rather than stripping out price variation from measured productivity to achieve consistency with our model, we could also have augmented the model to allow for imperfect competition and variable markups. This is the route taken by Bernard *et al.* (2003), who work with the assumption of Bertrand competition and limit pricing in which the lowest-cost supplier is constrained not to charge more than the second-lowest cost supplier. The drawback of this approach is that results are potentially sensitive to the particular choice of mechanism generating variation in mark-ups. In our context, data availability is an additional serious issue, since firm-level data or at least information about within-sector, across-firm productivity differences are required to implement the Bernard *et al.* (2003) methodology. Such data are not available in comparable form for the countries in our sample.

[24] Again, we restrict our sample to countries which are in our calibration sample and for which we have the necessary data. For manufacturing trade, we lose the same five countries as before due to aggregation (see footnote 21), plus Uzbekistan due to missing production data. Agricultural production data are unfortunately much less complete, restricting the estimation sample to 66 countries. Using all 78 countries for which production data is

Poisson QMLE can accomodate zero trade flows, which are common in our data, and leads to consistent parameter estimates even in the presence of heteroskedasticity in ε_{lj}. Appendix Table A.1 contains details of the estimation results, which are broadly in line with those from comparable specifications in the literature.[25] In robustness checks below, we will also use estimates of T_{lj} obtained by estimating a log-linearized version of (10) via OLS (see Appendix Table A.1 for results). The distance coefficients in both sets of estimations provide estimates for $(1-\sigma)\,\delta$ which, together with our estimated values of σ, yield estimates for δ and thus for $T_{lj} = dist_{lj}^{\delta_1} e^{\delta_2 d_{int,lj}}$.

We again use data on nominal and real expenditure on food and manufacturing goods from the ICP to obtain values for α and $\underline{A}$. In the model, the nominal expenditure share of manufacturing in GDP, and real consumption of agricultural goods per head, respectively, are given by:

$$S_{EMj} = \frac{E_{Mj}}{w_j L_j} = \alpha - \alpha \underline{A} \frac{P_{Aj}}{w_j}, \tag{11}$$

$$\frac{E_{Aj}}{P_{Aj} L_j} = (1-\alpha) w_j P_{Aj}^{-1} + \alpha \underline{A}. \tag{12}$$

Note that as $w_j \to \infty$, $S_{EMj} \to \alpha$; and as $w_j \to \underline{A} P_{Aj}$, $E_{Aj}/\,(P_{Aj} L_j) \to \underline{A}$. For our simulations below, we thus use the nominal expenditure share of manufacturing in total expenditure on food and manufacturing $(S_{EMj}/\,(S_{EMj} + S_{EAj}))$ of the richest country (Luxembourg) in our data as a proxy for α. Likewise, we use the real food expenditure per worker of the poorest country (Zambia) as a proxy for $\underline{A}$.[26]

4.2 Replication of Stylized Facts

We now present results for the same regressions as in Tables 1 and 2, but this time we use simulated rather than actual data for the year 2000.[27] That is, we use the calibrated model to generate artificial data on manufacturing shares for the developing countries in our simulation sample.[28] Note that our model also generates data for *per-capita* GDP (equal to wages in the model), GDP (wages times population size) and centrality (calculated according to (2), using the same distance data but replacing GNP with model-generated GDP). Thus, we use generated data for both dependent and independent variables in the regressions below, consistent with the notion that we would like to evaluate whether our model is comparable to the data generating

available only leads to minor changes in the estimates for the trade cost elasticities which have no impact on the following results (details available from the authors).

[25] In a recent meta study, Disdier and Head (2008) report that the mean distance elasticitiy of the 1,467 estimates they analyze is -0.9, very close to our Poisson estimates. Most studies exclude intranational trade but those that include it find estimates of comparable magnitude to ours. For example, Wei (1996) estimates $(1-\sigma_M)\,\delta_{M2} = -2.27$ for a sample of OECD countries between 1982 and 1994, compared to $(1-\sigma_M)\,\delta_{M2} = -1.99$ in our estimation.

[26] Again, also see Appendix B for further details on the ICP data. We have also experimented with using averages across the three or five richest/poorest countries, with similar results in the quantitative examples below. A significant downside of using more countries is, however, that the resulting higher estimates of $\underline{A}$ implied that we needed to drop countries from the data for which the subsistence condition of the model ($\underline{A} < \theta_{Al}$) was violated.

[27] As already discussed in Section 2 (footnote 13), availability of expenditure and price data from the ICP prevents us from generating data for earlier years. The availability of productivity and workforce data in agriculture and manufacturing also worsens as we go back in time, although not as dramatically as for the ICP data.

[28] Again, the model is simulated for all 107 countries (developed and developing) as developed countries do of course play a major role in determining manufacturing shares and centrality of developing countries.

process in the real world. Population size and productivity data are of course directly used as model parameters, and are identical to the data used in the regressions from Section 2.

Table 4 presents regression results using our generated data which yield a similar picture as our earlier results using actual data.[29] The coefficient on centrality is positive and significant in all specifications. Likewise, relative productivity is never significantly negative. Similar to Table 2, it has a positive impact on industrialization in columns 1 and 5, but loses its significance as soon as we control for agricultural productivity. Thus, we replicate the basic findings highlighted in the introduction and in Section 2.

Tables 5 and 6 report a number of robustness checks. We first demonstrate that augmenting the model by a third, non-tradable sector (which can be thought of as services) does not change our previous results. We now model the representative individual's preferences from country j as

$$U_j = \alpha \ln(M_j - \underline{M}) + \beta \ln(A_j - \underline{A}) + (1 - \alpha - \beta) \ln S_j,$$

where A_j and M_j are defined as before, and $S_j = s_j$ is the locally produced services good.[30] Similar to agricultural and manufacturing varieties, services are produced using only labor with linear production technology $s_l = \theta_{Sl} L_{Sl}$ (where θ_{Sl} is labor productivity in services in country l). In Appendix D, we provide a more detailed exposition of the model, the resulting equilibrium conditions and analytical examples comparable to Section 4. As we show there, allowing for a service sector in the model complicates the analysis somewhat but the qualitative results go through as before.

Regarding the calibration of this augmented model, note that since the third sector is non-tradable and non-differentiated, we only require new estimates for α, $\underline{A}$, β, and $\underline{M}$ (see Appendix D for the modified procedure for obtaining them). In Table 5, we present the same set of regressions results as in Table 4, this time using the calibrated version of the three-sector model to generate our synthetic dataset. The results are both qualitatively and quantitatively very similar to before. We conclude that allowing for an additional non-tradable sector does not change our previous conclusions and we work with the initial two-sector model for the rest of this paper.[31]

[29] Note that the set of countries is the slightly different in Tables 1, 2 and 4 because of different data requirements. For generating our artificial data, we require the same independent variables as in Tables 1 and 2, but also employment in agriculture and manufacturing to compute workforce sizes (L_j). On the other hand, we do not need data on manufacturing shares as before. Running regressions on actual and generated data for the 76 countries present in both samples yields very similar results to Tables 1, 2 and 4 (available from the authors).

[30] We also worked with an alternative specification of preferences as in Kongsamut, Rebelo and Xie (2001), where $U_j = \alpha \ln(M_j) + \beta \ln(A_j - \underline{A}) + (1 - \alpha - \beta) \ln (S_j + \underline{S})$. As long as $\underline{A} > \underline{M}$ in our original preference specification (which is the case empirically, see Table 3) this new specification leads to identical qualitative results. Quantitative results from the calibrated model are also very similar to before although the calibration of preference parameters is slightly more involved in this new case (results available from the authors).

[31] The three-sector model also allows for an interesting extension of our data and results. Since we are now modelling the services sector as well, our model should be better suited to model the sectoral structure of developed countries as well. According to our model, comparative advantage and centrality should also play a role in determining manufacturing shares for this group of countries (although the impact of non-homotheticity will be smaller at higher income levels). In Appendix Tables A.2 and A.3, we thus present results for the full set of countries for which we have data (both developing and developed). Table A.2 uses actual data, while Table A.3 uses the data just generated by our three sector model. Again, both sets of results are similar, confirming that the model also performs well when applied to all countries.

In Table 6, we report a number of additional robustness checks for the two-sector model. The first three columns use our alternative set of substitution elasticity estimates ($\sigma_M = 2.0$ and $\sigma_A = 2.6$). In columns 4-6, we use ordinary least squares to estimate equation (10), leading to alternative estimates for δ_{1M}, δ_{1A}, δ_{2M}, and δ_{2A}. Finally, in columns 7-9 we use producer prices rather than consumer prices to deflate relative productivities (see Appendix B). As shown, none of these changes alters the basic message from Table 4. Centrality is positive and significant throughout. Relative productivity is positive and significant when we do not control for agriculture productivity, and it is always insignificant when we do.[32]

4.3 Model Evaluation and Comparison

In Table 7 we compare our baseline calibration with subsistence consumption and positive but finite trade costs (see Tables 3 and 4) to three alternative calibrations. In the first, we set the trade cost parameters (δ_{1M}, δ_{1A}, δ_{2M}, and δ_{2A}) to infinity ('autarky'). In the second, we set the same parameters to zero ('free trade'), and in the third, we use our original trade cost estimates but eliminate non-homotheticity by setting $\underline{A}= 0$.

Our comparison uses two criteria. First, can the model qualitatively replicate the correlations found in the data between comparative advantage and centrality, on the one hand, and manufacturing shares, on the other hand? Second, how well do the original and the three alternative parameterizations do in terms of replicating actual manufacturing shares? To evaluate this second criterion, we regress actual on simulated manufacturing shares, and look at the sign and significance of the corresponding intercept and slope coefficient, as well as at the R^2 of the regression. With a perfect fit between actual and simulated data, we would expect an intercept equal to zero, and a slope coefficient and R^2 of unity. Note that in this simple regression, the R^2 is also equal to the squared correlation coefficient between actual and generated data which we also report for convenience.

In columns 1-3 of Table 7, we restate our baseline results from Table 4 and report the measures of fit described above. Our baseline parameterization not only reproduces our stylized facts but also produces a moderately good fit to the data (see the last lines of the table). The intercept of the regression of actual on simulated manufacturing shares is zero, the slope coefficient is 0.8 and statistically indistinguishable from unity, and the correlation between actual and generated data is around +40%.

How does this compare to our three alternative parameterizations? Looking at free trade first, we see that the model's performance in this case is dismal with respect to all criteria (columns 4-6). As expected, the coefficient on comparative advantage is negative and strongly significant, whereas the one on centrality is negative in column 5 and turns insignificant once we include all regressors in column 6. The intercept from the regression of actual on simulated manufacturing shares is positive and significant, the slope coefficient close to zero and statistically significantly different from one, and the corresponding R^2 close to zero. The model's performance

[32] For conciseness, we omit the specifications also including population and *per-capita* GDP. Results are again similar to those for our baseline calibration shown in Table 4 (available from the authors upon request).

with infinitely high trade costs is somewhat better, in the sense that it can replicate the facts related to relative productivity (column 7-9). However, the coefficient on centrality is insignificant throughout as expected.[33] Also, while the fit of the model to the data is much better under autarky than with free trade, it is outperformed by our baseline calibration on all our criteria (R^2, closeness of the regression intercept to zero and of the slope to unity). Finally, the performance of our model deteriorates substantially when we eliminate subsistence consumption. As expected, the model is no longer able to replicate our stylized facts and and the regression R^2 drops to just 0.04.[34]

4.4 Counterfactuals Experiments

Having demonstrated the empirical relevance of our calibrated model, we use it to perform a number of counterfactual simulations.

In the first counterfactual, we analyze the role of one of the key drivers of industrialization in existing models of structural change (agriculture productivity) in our setting with positive but finite trade cost. We start from the baseline calibration of our model and gradually increase agricultural productivity. We do so separately for each country and record the associated change in the manufacturing share in that country. We then compare the direction and magnitude of these changes across more and less remote developing countries, as measured by our centrality measure in (2). Figure 3 plots the increase in agricultural productivity on the horizontal axis, and the resulting average change in the manufacturing share for three groups of countries on the vertical axis: developing countries in the lowest tercile of our centrality measure, developing countries in the middle tercile and developing countries in the top tercile.[35] As seen, the less central an economy is, the more its manufacturing share increases following a given increase in agricultural productivity.[36] Interestingly, higher agricultural productivity on average raises manufacturing shares across all terciles of the centrality measure, with only a small minority of

[33]The fact that the coefficient on centrality is not exactly zero under autarky is of course due to functional form misspecification, given that the true data generating process in the model is more complicated than the simple log-linear relationship postulated in our regression tables throughout. For future research, it would be interesting to investigate whether using functional forms directly implied by the model have higher explanatory power in the actual data as well. We note, however, that this does not invalidate our earlier comparisons based on log-linear specifications as the issue of functional form misspecification applies to both actual and generated data. If the underlying data generating process were similar in both samples, we would expect the same log-linear approximation to yield similar results (as indeed it does).

[34]We have also experimented with using lower but non-zero levels of subsistence consumption. As expected, the univariate correlation between relative productivity and manufacturing shares gradually declines as we lower $\underline{A}$ but it only becomes insignficant at below 10% of its original value. Likewise, the comparative advantage term in the specification controlling for agricultural productivity becomes negative and significant at values of around 25% of the original value.

[35]Figure 3 also plots 95% confidence intervals for the mean of each group, indicating that the means of the top and bottom tercile are statistically different from each other (we omit the confidence interval for the middle tercile to avoid clutter but note that it overlaps with the confidence intervals of the other terciles).

[36]Note that the effect of agricultural productivity increases does also depend on a country's initial per-capita income (pro-industrializing effects are smaller for richer countries). Thus, we also ran multivariate regressions of our simulated changes in manufacturing shares on centrality, controlling for wages (or, alternatively, for initial agricultural and manufacturing productivity) as well as population size. Centrality continued to have a positive and significant impact on simulated share changes.

countries experiencing a (very slight) decline. Put differently, what we called the "relative-demand effect" of agricultural productivity shocks outweighs their "comparative-advantage effect" for our baseline parameter estimates. While our model is of course highly stylized, this suggests that the de-industrialising effects of high agricultural productivity highlighted in parts of the literature (e.g., Matsuyama (1992, 2009)) might not be relevant in practice.

Our first experiment holds trade costs constant while changing the comparative advantage of countries. Our second and third experiments do the opposite. We gradually increase openness to trade in our model and observe how existing comparative advantage patterns manifest themselves. We do so by gradually eliminating the home bias component of our trade cost matrix, again starting from the baseline calibration of our model (i.e., we gradually lower δ_{2M} and δ_{2A} and recompute the trade cost matrix T). The home bias parameters can be thought of as a summary measure of all trade barriers which render international trade more costly relative to intranational trade. As such, it is the closest equivalent in our model to policy-induced changes such a regional trade agreements and multilateral trade negotiations in the WTO which try to eliminate both tariff and non-tariff barriers to trade.[37] In our second experiment, we reduce trade barriers between all countries (i.e., we use the lower values for δ_{2M} and δ_{2A} for all elements of our trade cost matrix). In our third experiment, we only reduce trade costs between countries in the same geographic region, using the World Bank's classification of countries into seven broad regions.[38]

Figure 4 shows results for the global reduction in trade barriers. We again plot the resulting changes in manufacturing shares for three groups of countries, but this time group countries into terciles according to their comparative advantage in agriculture (as measured by the parameter ratio θ_{Al}/θ_{Ml}). As expected, countries increasingly specialize according to their comparative advantage as the degree of openness increases.[39] We also note that the average impact of trade liberalization on manufacturing activity is positive and that the majority of developing countries see an increase in manufacturing shares. This effect arises because freer trade increases real income levels and thus shifts expenditure towards manufacturing.

Figure 5 shows the results for a regional reduction in trade barriers. The picture looks broadly similar to before, but with two important exceptions. First, the average increase in manufacturing shares is smaller. This is simply due to the fact that the impact of regional trade liberalization on real wages is lower than that of the full liberalization, and the average increase in manufacturing

[37] Lowering the elasticity of trade costs with respect to distance (δ_{1M} and δ_{1A}) as well yields qualitatively similar results. We focus on the home bias component here because it corresponds more closely to policy-induced barriers and because there is not much evidence that trade cost elasticities have become smaller in absolute terms over the past decades (see Disdier and Head, 2008, although we note that their evidence relates to the distance elasticity of trade flows, i.e., the compositive parameter $(1 - \sigma_M)\,\delta_{1M}$).

[38] The seven regions are East Asia & Pacific, Europe & Central Asia, Latin America & Carribean, North America, Middle East & North Africa, South Asia, and Sub-Saharan Africa. Of these, all but North America contain developing countries according to our definition (see Appendix A). We also note that existing regional trade agreements broadly follow this classification (some prominent examples include the European Union, Mercosur, NAFTA and ASEAN).

[39] We again omit the 95% confidence interval for the mean of the middle tercile but note that this time, none of the three confidence intervals overlaps with each other. As before, multivariate regressions of our simulated changes in manufacturing shares on comparative advantage and additional controls (centrality, population size) yield the same results.

expenditure thus smaller. Secondly, the induced specialization is more muted than in the case of a global reduction in trade barriers. Note that this is not due to the fact that some trade barriers still remain in this case.[40] Rather, it is because differences in comparative advantage (as measured by the productivity ratio θ_{Al}/θ_{Ml}) are less pronounced across neighbouring countries. At least for the case of free trade, the underlying intuition is easily seen from the analytical results in Section 3.2.2. What matters for the variation in manufacturing shares across countries in our model is the variation in relative productivity levels – and this variation is smaller within than across regions. For example, the standard deviation of our comparative advantage measures across developing countries is 0.36 globally but only 0.23 on average if we look region by region. Across all 107 developed and developing countries used in our calibration, it is 0.34 globally but only 0.20 regionally. More generally, bilateral differences in comparative advantage seem to increase significantly with bilateral distance, at least in our data.[41]

Again, while our model is highly stylized, we think that this differential impact of global and regional reductions in trade barriers on industrialization adds an interesting element to the ongoing discussion on the relative merits of regional trade agreements vs multilateral liberalization in the WTO.

5 Conclusion

In this paper, we have drawn attention to two cross-sectional facts which, taken together, are not easily explained by existing models of industrialization. Drawing on a number of theoretical mechanisms well-known in the trade and economic geography literature, we presented a multi-location model with trade costs and non-homothetic preferences. Using analytical examples and a full-scale many-country calibration, we showed that this model can account for our stylized facts. Introducing trade cost and non-homotheticity also improved the fit of the model to cross-sectional data on manufacturing shares as compared to models with zero or infinite trade costs, or without non-homotheticity. Finally, the parameterized version of our model lends itself to counterfactual analysis which we illustrated through examples regarding agricultural productivity growth and a global vs regional reduction in trade barriers.

[40] For example, we also conducted counterfactual experiments in which we only reduced trade barriers between developed and developing countries, or among developing countries only. We found that in some scenarios, the average decline in manufacturing shares for the group of developing countries with the highest relative productivity in agriculture was actually more pronounced than in the case of a global reduction in trade barriers.

[41] To show this, we ran regressions of the form $\Delta pr_{lj} = a + \beta \times ldist_{lj} + e_{lj}$, where a is an intercept, $ldist_{lj}$ is the log of the bilateral distance between l and j and where $\Delta pr_{lj} = 1/relpr_{lj}$ if $relpr_{lj} < 1$ and $relpr_{lj}$ otherwise ($relpr_{ljt} = (\theta_{Al}/\theta_{Ml})/(\theta_{Aj}/\theta_{Mj})$); this definition of Δpr_{lj} assures that it takes on a value of unity for countries with identical relative productivities and is larger otherwise). A simple OLS regression yielded a coefficient on $ldist_{lj}$ of around 0.3 which was significant at the 1%-level. This result is robust to the inclusion of dummies for reporter and/or partner countries.

References

[1] Behrens, K., A. Lamorgese, G. Ottaviano, and T. Tabuchi (2009): "Beyond the Home Market Effect: Market Size and Specialization in a Multi-Country World," Journal of International Economics, vol. 79(2), 259-265.

[2] Behrens, K., G. Mion, Y. Murata and J. Südekum (2009): "Trade, Wages, and Productivity," CEPR DP 7369.

[3] Bernard, A.B., J. Eaton, J.B. Jensen, and S. Kortum (2003): "Plants and Productivity in International Trade," American Economic Review, 93(4), 1268-1290.

[4] Breinlich, H. (2006): "The Spatial Income Structure in the European Union – What Role for Economic Geography?," Journal of Economic Geography, 6(5), 593-617.

[5] Broda, C., J. Greenfield, and D. Weinstein (2006): "From Groundnuts to Globalization: A Structural Estimate of Trade and Growth," NBER Working Paper 13041.

[6] Broda, C. and D. Weinstein (2006): "Globalization and the Gains from Variety," Quarterly Journal of Economics, 121 (2), 541-585.

[7] Caselli, F. and Coleman, W.J. (2001): "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," Journal of Political Economy, 109, 584-616.

[8] Coleman, W.J. (2007): "Accomodating Emerging Giants," mimeo, Duke University.

[9] Crafts, N. (1980): "Income Elasticities of Demand and the Release of Labor by Agriculture during the British Industrial Revolution," Journal of European Economic History, 9, 153-168.

[10] Davis, D. and D. Weinstein (2003): "Market Access, Economic Geography, and Comparative Advantage: An Empirical Test," Journal of International Economics, 59(1), 1-23.

[11] Desmet, K. and E. Rossi-Hansberg (2011): "Spatial Development," mimeo, Princeton.

[12] Disdier, A.-C. and K. Head (2008): "The Puzzling Persistence of the Distance Effect on Bilateral Trade," Review of Economics and Statistics 90, 37-41.

[13] Duarte, M. and D. Restuccia (2010), "The Role of the Structural Transformation in Aggregate Productivity," Quarterly Journal of Economics, 129-173.

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

[15] Eaton, J., S. Kortum and F. Kramarz (2011): "An Anatomy of International Trade: Evidence from French Firms," Econometrica, 79(5), 1453-1498.

[16] Echevarria, C. (1997): "Changes in Sectoral Composition Associated with Economic Growth," International Economic Review, 38(2), 431-452.

[17] Epifani, P. and G. Gancia (2006): "Increasing Returns, Imperfect Competition, and Factor Prices," Review of Economics and Statistics, 88(4), 583-598.

[18] Fajgelbaum, P.D., G.M. Grossman and E. Helpman (2011): "Income Distribution, Product Quality, and International Trade," Journal of Political Economy, 119, 721-765.

[19] Feenstra, R. (1994): "New Product Varieties and the Measurement of International Prices," American Economic Review, 84(1), 157-177.

[20] Feenstra, R., J. Romalis, and P. Schott (2002): "U.S. Imports, Exports and Tariff Data, 1989-2001," mimeo, University of California, Davis.

[21] Feenstra, R., R. Lipsey, H. Deng, A. Ma, and H. Mo (2005): "World Trade Flows, 1962-2000," NBER Working Paper 11040.

[22] Fieler, A.C. (2011): "Non-homotheticity and Bilateral Trade: Evidence and a Quantitative Explanation," Econometrica, 79(4), 1069-1101.

[23] Galor, O. and A. Mountford (2008), "Trading Population for Productivity: Theory and Evidence," Review of Economic Studies, 75, 1143–1179.

[24] Golub, Stephen S. and Hsieh, Chang-Tai (2000): "Classical Ricardian Theory of Comparative Advantage Revisited," Review of International Economics, 8(2), 221-234.

[25] Hanson, G. and C. Xiang (2004): "The Home Market Effect and Bilateral Trade Patterns," American Economic Review, 94 (4), 1108-1129.

[26] Harris, J. (1954): "The market as a factor in the localization of industry in the United States" Annals of the Association of American Geographers, 64, 315-348.

[27] Head, K. and T. Mayer (2006): "Regional Wage and Employment Responses to Market Potential in the EU," Regional Science and Urban Economics 36(5), 573-595.

[28] Kongsamut, P., S. Rebelo and D. Xie (2001): "Beyond Balanced Growth," Review of Economic Studies, 68, 869-882.

[29] Krugman, Paul R. (1980): "Scale Economies, Product Differentiation, and the Pattern of Trade," American Economic Review, 70, 950-959.

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

[31] Krugman, P. and E. Helpman (1985): "Market Structure and Foreign Trade," MIT Press, Cambridge, MA.

[32] Linder, S. (1961): An Essay on Trade and Transformation. Uppsala: Almqvist & Wicksells.

[33] Markusen, J. (1986): “Explaining the Volume of Trade: An Eclectic Approach," American Economic Review, 76(5), 1002-1011.

[34] Matsuyama, K. (1992): “Agricultural Productivity, Comparative Advantage, and Economic Growth,” Journal of Economic Theory, 58, 317-334.

[35] Matsuyama, K. (2000): “A Ricardian Model with a Continuum of Goods Under Nonhomothetic Preferences: Demand Complementarities, Income Distribution and North-South Trade,” Journal of Political Economy, 108, 1093-1120.

[36] Matsuyama, K. (2009): “Structural Change in an Interdependent World: A Global View of Manufacturing Decline,” Journal of the European Economic Association P&P.

[37] Melitz, M.J. and G.I.P. Ottaviano (2008): “Market Size, Trade, and Productivity,” Review of Economic Studies, 75, 295-316.

[38] Mitra, Devashish and V. Trindade (2005): “Inequality and Trade,” Canadian Journal of Economics, 38(4), 1253-1271.

[39] Murata, Y. (2008): “Engel’s Law, Petty’s Law, and Agglomeration,” Journal of Development Economics, 87, 161-177.

[40] Murphy, Kevin M., A. Shleifer and R.V. Vishny (1989a): “Industrialization and the Big Push,” Journal of Political Economy, 97(5), 1003-1026.

[41] Murphy, Kevin M., A. Shleifer and R.V. Vishny (1989b): “Income Distribution, Market Size, and Industrialization,” Quarterly Journal of Economics, 104(3), 537-564.

[42] Ngai, L.R. and C.A. Pissarides (2007): “Structural Change in a Multisector Model of Growth,” American Economic Review, 97, 429-443.

[43] Puga, D. and A.J. Venables (1999): “Agglomeration and Economic Development: Import Substitution vs. Trade Liberalisation,” Economic Journal, 109, 292-311.

[44] Redding, S. and A.J. Venables (2004): “Economic Geography and International Inequality,” Journal of International Economics, 62, 53-82.

[45] Santos-Silva, J. and S. Tenreyro (2005): “The Log of Gravity,” Review of Economics and Statistics, 88, 641-658.

[46] Syrquin, M. and H. Chenery (1989): “Patterns of Development, 1950 to 1983”, World Bank Discussion Paper 41.

[47] Wei, S.J. (1996): “Intra-National versus International Trade: How Stubborn are Nations in Global Integration?,” NBER Working Paper 5531.

[48] Yi, K.M. and J. Zhang (2010): “Structural Change in an Open Economy,” mimeo, Federal Reserve Bank of Philadelphia.

A Appendix A: Country Lists and Data used in Cross-Country Regressions

Country List I - Cross-Country Regressions. Albania (ALB); Algeria (DZA); Angola (AGO); Argentina (ARG); Armenia (ARM); Azerbaijan (AZE); Bangladesh (BGD); Barbados (BRB); Belarus (BLR); Belize (BLZ); Benin (BEN); Bhutan (BTN); Bolivia (BOL); Botswana (BWA); Brazil (BRA); Bulgaria (BGR); Burkina Faso (BFA); Burundi (BDI); Cambodia (KHM); Cameroon (CMR); Cape Verde (CPV); Central African Republic (CAF); Chad (TCD); Chile (CHL); China (CHN); Colombia (COL); Comoros (COM); Congo, Dem. Rep. (ZAR); Congo, Rep. (COG); Costa Rica (CRI); Cote d'Ivoire (CIV); Croatia (HRV); Czech Republic (CZE); Dominican Republic (DOM); Ecuador (ECU); Egypt, Arab Rep. (EGY); El Salvador (SLV); Equatorial Guinea (GNQ); Eritrea (ERI); Estonia (EST); Ethiopia (ETH); Fiji (FJI); Gabon (GAB); Gambia, The (GMB); Georgia (GEO); Ghana (GHA); Grenada (GRD); Guatemala (GTM); Guinea (GIN); Guinea-Bissau (GNB); Guyana (GUY); Honduras (HND); Hungary (HUN); India (IND); Indonesia (IDM); Iran, Islamic Rep. (IRN); Jamaica (JAM); Jordan (JOR); Kazakhstan (KAZ); Kenya (KEN); Korea, Rep. (KOR); Kyrgyz Republic (KGZ); Lao PDR (LAO); Latvia (LVA); Lesotho (LSO); Lithuania (LTU); Macedonia, FYR (MKD); Madagascar (MDG); Malawi (MWI); Malaysia (MYS); Mali (MLI); Mauritania (MRT); Mexico (MEX); Moldova (MDA); Mongolia (MNG); Morocco (MAR); Mozambique (MOZ); Namibia (NAM); Nepal (NPL); Nicaragua (NIC); Niger (NER); Nigeria (NGA); Oman (OMN); Pakistan (PAK); Panama (PAN); Papua New Guinea (PNG); Paraguay (PRY); Peru (PER); Philippines (PHL); Poland (POL); Romania (ROM); Rwanda (RWA); Saudi Arabia (SAU); Senegal (SEN); Sierra Leone (SLE); Slovak Republic (SVK); South Africa (ZAF); Sri Lanka (LKA); St. Lucia (LCA); St. Vincent and the Grenadines (VCT); Sudan (SDN); Suriname (SUR); Syrian Arab Republic (SYR); Tanzania (TZA); Thailand (THA); Togo (TGO); Trinidad and Tobago (TTO); Tunisia (TUN); Turkey (TUR); Uganda (UGA); Ukraine (UKR); Uruguay (URY); Uzbekistan (UZB); Venezuela, RB (VEN); Vietnam (VNM); Yemen, Rep. (YEM); Zambia (ZMB); Zimbabwe (ZWE).

Country List II - Calibration Sample. Developing countries: Albania (ALB); Algeria (DZA); Argentina (ARG); Armenia (ARM); Azerbaijan (AZE); Bangladesh (BGD); Barbados (BRB); Belize (BLZ); Bolivia (BOL); Botswana (BWA); Brazil (BRA); Bulgaria (BGR); Cambodia (KHM); Cameroon (CMR); Chile (CHL); China (CHN); Colombia (COL); Costa Rica (CRI); Croatia (HRV); Czech Republic (CZE); Dominican Republic (DOM); Ecuador (ECU); Egypt, Arab Rep. (EGY); El Salvador (SLV); Estonia (EST); Georgia (GEO); Ghana (GHA); Guatemala (GTM); Guyana (GUY); Honduras (HND); Hungary (HUN); Indonesia (IDM); Jamaica (JAM); Jordan (JOR); Kazakhstan (KAZ); Korea, Rep. (KOR); Kyrgyz Republic (KGZ); Latvia (LVA); Lesotho (LSO); Lithuania (LTU); Macedonia, FYR (MKD); Malaysia (MYS); Mexico (MEX); Moldova (MDA); Mongolia (MNG); Morocco (MAR); Namibia (NAM); Nepal (NPL); Nicaragua (NIC); Oman (OMN); Pakistan (PAK); Panama (PAN); Papua New Guinea (PNG); Paraguay (PRY); Peru (PER); Philippines (PHL); Poland (POL); Romania (ROM); Russian Federation (RUS); Saudi Arabia (SAU); Slovak Republic (SVK); South Africa (ZAF); Sri Lanka (LKA); St. Lucia (LCA); St. Vincent and the Grenadines (VCT); Suriname (SUR); Syrian Arab Republic (SYR); Tanzania (TZA); Thailand (THA); Trinidad and Tobago (TTO); Turkey (TUR); Ukraine (UKR); Uruguay (URY); Uzbekistan (UZB); Venezuela, RB (VEN); Vietnam (VNM); Yemen, Rep. (YEM); Zambia (ZMB); Zimbabwe (ZWE).

Developed countries: Australia (AUS); Austria (AUT); Belgium (BEL); Brunei Darussalam (BRN); Canada (CAN); Cyprus (CYP); Denmark (DKF); Finland (FIN); France (FRA);

Germany (DEU); Greece (GRC); Iceland (ISL); Ireland (IRL); Italy (ITA); Japan (JPN); Luxembourg (LUX); Netherlands (NLD); New Zealand (NZL); Norway (NOR); Portugal (PRT); Singapore (SGP); Slovenia (SVN); Spain (ESP); Sweden (SWE); Switzerland (CHE); United Arab Emirates (ARE); United Kingdom (GBR); United States (USA).

Data used in Cross-Country Regressions.

- Share of manufacturing value added in GDP: World Development Indicators (World Bank) and national statistical offices.
- Value added per worker in agriculture and manufacturing (in 2000 USD): World Development Indicators, United Nations Industrial Statistics Database (UNIDO), and national statistical offices.
- GDP, GNP and *per-capita* GDP (2000 USD): World Development Indicators.
- Population size and size of workforce in agriculture and manufacturing: World Development Indicators.
- Bilateral distances between countries: CEPII Bilateral Distances Database.
- Sector-specific PPP exchange rates: International Comparison Project (ICP).

B Appendix B: Using ICP Data to Proxy for Prices

In Sections 2 and 4 we use data from the International Comparison Project (ICP) to strip out price variation from measured productivity. To understand this approach, note that the ICP provides data on a number of expenditure categories in both current U.S. dollars and so-called international dollars ($I). One $I is the amount of goods and services one U.S. dollar would purchase in the USA in the base period (2005 in our case as no data were available for 2000). Converting expenditure from current U.S. dollars into $I thus removes any price differences across countries and basically converts expenditures into quantities using implicit aggregators. By comparing local expenditures in U.S. dollars and international dollars, one can derive country-product-specific PPP exchange rates which capture price differences across country. For example, *per-capita* expenditure on food in current U.S. dollars was $2,040 in 2005 in the United Kingdom but only 1,586 $I, yielding an implicit price of 1.29 (the price in the USA is normalized to 1). Dividing measured productivity levels ($p_{Ml}m_l/L_{Ml}$ and $p_{Al}a_l/L_{Al}$) by this price converts them into quantities per unit of labor used and thus into appropriate proxies for $\theta_{Ml} = m_l/L_{Ml}$ and $\theta_{Al} = a_l/L_{Al}$. We note that Echevarria (1997) uses a similar procedure, calculating proxies for agricultural and manufacturing prices by dividing expenditures in U.S. dollars by expenditures in international dollars.

One problem with the above approach is that we are implicitly using consumer prices rather than producer prices to deflate production. In terms of our model, ICP prices are proxies for P_{Ml} and P_{Al}, not p_{Ml} and p_{Al}. As a robustness check in Section 4, we therefore use the definition of P_{Ml} and P_{Al} to extract information on p_{Ml} and p_{Al} in a model-consistent way. In our model,

$$P_{Mj} = \left[\sum_{l=1}^{R} \left(p_{Ml}T_{lj}^{M}\right)^{1-\sigma_M}\right]^{\frac{1}{1-\sigma_M}} \tag{13}$$

$$P_{Aj} = \left[\sum_{l=1}^{R} \left(p_{Al}T_{lj}^{A}\right)^{1-\sigma_A}\right]^{\frac{1}{1-\sigma_A}}, \tag{14}$$

Together with data on the elasticities of substitution and trade costs which we have obtained independently as part of our calibration strategy, we can solve the above system of equations for p_{Ml} and p_{Al}. In practice, consumer and implied producer prices are almost identical, with a correlation coefficient of above 99% and a level difference of on average less than 4% for manufacturing and less than 1% for agriculture.

C Appendix C: Analytical Examples

This appendix provides the analytical treatment of the examples discussed in Section 3.2.3.

Example 1 Consider a three-country world, $R = 3$, and a geographic structure such that country 1 takes a "central" position while countries 2 and 3, which are fully symmetric, are in the "periphery": we model this by assuming that country 1 can trade with both 2 and 3 at positive but finite trade costs ($T_{12} = T_{21} = T_{13} = T_{31} = T > 1$) and that countries 2 and 3 cannot trade with one another ($T_{23} = T_{32} = \infty$).[42] Trade costs are assumed equal across sectors. We simplify further by assuming $\sigma_M = \sigma_A = \sigma$. Finally, we choose all parameters to be identical across countries (except for the bilateral trade costs) and, in particular, we set $\theta_{Aj} = \theta_{Mj} = L_j = 1$. Profiting from the symmetry we have imposed, let us normalize $w_2 = w_3 = 1$.

It is easy to show that we cannot have an equilibrium in which $w_1 = 1$, as the model's market clearing conditions would be violated. We can prove this by contradiction. If it were the case that $w_1 = w_2 = w_3 = 1$, then aggregate labor demand would be different across countries:

$$L_{M1} + L_{A1} = \frac{1}{2T^{1-\sigma}+1} + \frac{2}{T^{\sigma-1}+1} > L_{M2} + L_{A2} = \frac{1}{2+T^{\sigma-1}} + \frac{1}{1+T^{1-\sigma}}. \tag{15}$$

Thus, it must be the case that $w_1 > w_2 = w_3$. Due to the non-homotheticity of preferences, this implies that country 1's expenditure is biased towards manufacturing: $E_{M1} > E_{M2}$. As discussed above, positive trade costs lead this bias in demand for manufacturing goods to favor country 1's manufacturing industry primarily:

$$L_{M1} = \frac{1}{w_1^{1-\sigma} + 2T^{1-\sigma}} E_{M1} + \frac{2}{w_1^{1-\sigma} + T^{\sigma-1}} E_{M2}, \tag{16}$$

$$L_{M2} = \frac{1}{w_1^{1-\sigma} T^{\sigma-1} + 2} E_{M1} + \frac{1}{(w_1 T)^{1-\sigma} + 1} E_{M2}. \tag{17}$$

Establishing analytical results here is difficult, but the condition $2 > T^{\sigma-1}$, for example, is sufficient for $L_{M1} > L_{M2}$, which implies a larger manufacturing share in the central country.

Example 2 Again assume $R = 3$ and that all parameters are identical across countries (except for the bilateral trade costs) and, in particular, that $\theta_{Aj} = \theta_{Mj} = L_j = 1$, $\sigma_A = \infty$, and $\sigma_M > 1$ but finite. Again, we consider a geographic structure such that country 1 takes a "central" position while countries 2 and 3 are in the "periphery": here we model this by assuming that country 1 can trade freely with both 2 and 3 ($T_{12} = T_{21} = T_{13} = T_{31} = 1$) and that countries 2 and 3 cannot trade with one another ($T_{23} = T_{32} = \infty$).[43] Trade costs are again equal across sectors. We take the agricultural good as the numéraire. Under incomplete specialization for all

[42] For the sake of the argument, we rule out the possibility that countries 2 and 3 can trade via country 1.

[43] We again rule out the possibility that countries 2 and 3 can trade via country 1.

countries, the labor market equilibrium conditions comprise equations

$$L_{M1} = \frac{1}{3}\alpha\left(1-\underline{A}\right)+\alpha\left(1-\underline{A}\right)=\frac{4}{3}\alpha\left(1-\underline{A}\right), \tag{18}$$

$$L_{M2} = L_{M3}=\frac{1}{3}\alpha\left(1-\underline{A}\right)+\frac{\alpha}{2}\left(1-\underline{A}\right)=\frac{5}{6}\alpha\left(1-\underline{A}\right). \tag{19}$$

It is easy to show that in this case country 1's manufacturing share is larger than that of countries 2 and 3, since $L_{M1} > L_{M2} = L_{M3}$. If parameter values in this incomplete specialization scenario yielded $L_{M1} > 1$, then country 1 would specialize completely in manufacturing.[44] In this case, the labor market equilibrium conditions comprise equations

$$L_{M1} = w_1^{-\sigma_M}\left[\frac{\alpha\left(w_1-\underline{A}\right)}{\left(2+w_1^{1-\sigma_M}\right)}+\frac{2\alpha\left(1-\underline{A}\right)}{\left(1+w_1^{1-\sigma_M}\right)}\right]=1, \tag{20}$$

$$L_{M2} = L_{M3}=\frac{\alpha\left(w_1-\underline{A}\right)}{\left(2+w_1^{1-\sigma_M}\right)}+\frac{\alpha\left(1-\underline{A}\right)}{\left(1+w_1^{1-\sigma_M}\right)}<1, \tag{21}$$

which imply $w_1 > w_2 = w_3 = 1$. Notice that the mechanism discussed in this example does not depend on the non-homotheticity of preferences: assuming $\underline{A}= 0$ would not change the result here.[45]

Example 3 Consider many countries (R large). For simplicity, we assume again $\sigma_M = \sigma_A = \sigma$. All country-pairs face the same bilateral trade costs: $T_{jl}^M = T_{jl}^A = T > 1$ for all $j \neq l$. All countries have the same population size and productivities, $\theta_{Aj} = \theta_{Mj} = L_j = 1$ for all j, except for $\theta_{A1} > 1$. By symmetry, we can normalize $w_j = 1$ for all $j \neq l$. From the model's equilibrium conditions,

$$\frac{L_{M1}}{L_{A1}} = \frac{P_{M1}^{\sigma-1}E_{M1}+\sum_{l\neq 1}T^{1-\sigma}P_{Ml}^{\sigma-1}E_{Ml}}{\theta_{A1}^{\sigma-1}\left[P_{A1}^{\sigma-1}E_{A1}+\sum_{l\neq 1}T^{1-\sigma}P_{Al}^{\sigma-1}E_{Al}\right]}, \tag{22}$$

$$\frac{L_{Mj}}{L_{Aj}} = \frac{P_{Mj}^{\sigma-1}E_{Mj}+\sum_{l\neq j}T^{1-\sigma}P_{Ml}^{\sigma-1}E_{Ml}}{P_{Aj}^{\sigma-1}E_{Aj}+\sum_{l\neq j}T^{1-\sigma}P_{Al}^{\sigma-1}E_{Al}}, \tag{23}$$

for countries 1 and j. Assuming that trade costs are such that countries consume sizable amounts of foreign goods, one can neglect the effect of θ_{A1} on the price levels P_{Ml} and P_{Al}. A high θ_{A1} therefore has a direct effect in the denominator of equation (22) and an indirect effect via a high w_1 in the terms E_{M1} and E_{A1} of both equations. Notice first that the direct effect of θ_{A1} raises country 1's agricultural share in GDP (the comparative-advantage effect). Second, a higher w_1 (due to a higher θ_{A1}) tilts relative expenditure towards manufacturing in both country 1 and country j because of the non-homotheticity in demand, but more so in country 1 due to the presence of trade costs. As discussed above, this relative-demand effect operates in the direction opposite to the comparative-advantage effect.

[44] Under the assumption $\sigma_A = \infty$, there is no need for every country to produce its own "variety" of the agricultural good.

[45] A third mechanism which could generate higher levels of industrialization in the center is based on the manufacturing industry having access to both a constant returns to scale and an increasing returns to scale (IRS) production technique (see Murphy et al. (1989a) and (1989b)). In this case, central locations would be the first, *ceteris paribus*, to reach the critical level of demand that makes IRS production profitable. This mechanism is absent from our model, as we assume constant returns to scale across sectors.

D Appendix D: A Three-Sector Model

This appendix works out the three-sector model of Section 4.2, where the third sector, services, is assumed to be nontraded. We allow for non-homotheticities in demand to affect the manufacturing sector, too, as this is has been considered in the literature relatively often (see, for example, Matsuyama (2009)).

D.1 Demand Side

The individual's preferences are now

$$U_j = \alpha \ln(M_j - \underline{M}) + \beta \ln(A_j - \underline{A}) + (1 - \alpha - \beta) \ln S_j, \tag{24}$$

with

$$\begin{aligned} A_j &= \left[\sum_{l=1}^{R} a_{lj}^{(\sigma_A - 1)/\sigma_A}\right]^{\frac{\sigma_A}{\sigma_A - 1}} & (25) \\ M_j &= \left[\sum_{l=1}^{R} m_{lj}^{(\sigma_M - 1)/\sigma_M}\right]^{\frac{\sigma_M}{\sigma_M - 1}}, & (26) \\ S_j &= s_j, & (27) \end{aligned}$$

where $\alpha, \beta, \alpha + \beta \in (0, 1)$, $\underline{A} < \theta_{Al}$, $\underline{M} < \theta_{Ml}$, $\sigma_A, \sigma_M > 1$. The individual's budget constraint is

$$P_{Aj} A_j + P_{Mj} M_j + P_{Sj} S_j = w_j. \tag{28}$$

As we discuss below, total income equals labor income, as profits are zero. The price indices in the budget constraint are

$$\begin{aligned} P_{Aj} &= \left(\sum_{l=1}^{R} p_{Alj}^{1-\sigma_A}\right)^{\frac{1}{1-\sigma_A}} = \left[\sum_{l=1}^{R} \left(p_{Al} T_{lj}^{A}\right)^{1-\sigma_A}\right]^{\frac{1}{1-\sigma_A}}, & (29) \\ P_{Mj} &= \left(\sum_{l=1}^{R} p_{Mlj}^{1-\sigma_M}\right)^{\frac{1}{1-\sigma_M}} = \left[\sum_{l=1}^{R} \left(p_{Ml} T_{lj}^{M}\right)^{1-\sigma_M}\right]^{\frac{1}{1-\sigma_M}}, & (30) \\ P_{Sj} &= p_{Sj}. & (31) \end{aligned}$$

where $T_{lj}^{A}, T_{lj}^{M} \geq 1$, $T_{jj}^{A}, T_{jj}^{M} = 1$. Implicit here is the assumption that sector S is non-traded.

Aggregating across all individuals/countries yields the following demands for varieties (net of trade costs):

$$\begin{aligned} a_l &= p_{Al}^{-\sigma_A} \sum_{j=1}^{R} \left(T_{lj}^{A}\right)^{-\sigma_A} P_{Aj}^{\sigma_A - 1} E_{Aj}, & (32) \\ m_l &= p_{Ml}^{-\sigma_M} \sum_{j=1}^{R} \left(T_{lj}^{M}\right)^{-\sigma_M} P_{Mj}^{\sigma_M - 1} E_{Mj}, & (33) \\ s_l &= p_{Sl}^{-1} E_{Sl}, & (34) \end{aligned}$$

where

$$
\begin{aligned}
E_{Aj} &= \left[\underline{A}P_{Aj} + \beta\left(w_j - \underline{A}P_{Aj} - \underline{M}P_{Mj}\right)\right]L_j, &(35)\\
E_{Mj} &= \left[\underline{M}P_{Mj} + \alpha\left(w_j - \underline{A}P_{Aj} - \underline{M}P_{Mj}\right)\right]L_j, &(36)\\
E_{Sj} &= \left[(1-\alpha-\beta)\left(w_j - \underline{A}P_{Aj} - \underline{M}P_{Mj}\right)\right]L_j. &(37)
\end{aligned}
$$

D.2 Production

Goods are produced with linear technologies:

$$
\begin{aligned}
a_l &= \theta_{Al}L_{Al}, &(38)\\
m_l &= \theta_{Ml}L_{Ml}, &(39)\\
s_l &= \theta_{Sl}L_{Sl}. &(40)
\end{aligned}
$$

Perfect competition implies

$$p_{.l} = \frac{w_l}{\theta_{.l}}. \tag{41}$$

D.3 Equilibrium

Equilibrium in the goods markets yields

$$
\begin{aligned}
a_l &= \theta_{Al}^{\sigma_A} w_l^{-\sigma_A}\left[\sum_{j=1}^{R}\left(T_{lj}^{A}\right)^{1-\sigma_A} P_{Aj}^{\sigma_A-1}E_{Aj}\right], &(42)\\
m_l &= \theta_{Ml}^{\sigma_M} w_l^{-\sigma_M}\left[\sum_{j=1}^{R}\left(T_{lj}^{M}\right)^{1-\sigma_M} P_{Mj}^{\sigma_M-1}E_{Mj}\right], &(43)\\
s_l &= \frac{\theta_{Sl}}{w_l}E_{Sl}. &(44)
\end{aligned}
$$

Labor demand:

$$
\begin{aligned}
L_{Al} &= \theta_{Al}^{\sigma_A-1} w_l^{-\sigma_A}\left[\sum_{j=1}^{R}\left(T_{lj}^{A}\right)^{1-\sigma_A} P_{Aj}^{\sigma_A-1}E_{Aj}\right], &(45)\\
L_{Ml} &= \theta_{Ml}^{\sigma_M-1} w_l^{-\sigma_M}\left[\sum_{j=1}^{R}\left(T_{lj}^{M}\right)^{1-\sigma_M} P_{Mj}^{\sigma_M-1}E_{Mj}\right], &(46)\\
L_{Sl} &= \frac{E_{Sl}}{w_l}. &(47)
\end{aligned}
$$

Full employment requires

$$L_{Al} + L_{Ml} + L_{Sl} = L_l \tag{48}$$

or

$$\theta_{Al}^{\sigma_A-1} w_l^{-\sigma_A}\left[\sum_{j=1}^{R}\left(T_{lj}^{A}\right)^{1-\sigma_A} P_{Aj}^{\sigma_A-1}E_{Aj}\right] + \theta_{Ml}^{\sigma_M-1} w_l^{-\sigma_M}\left[\sum_{j=1}^{R}\left(T_{lj}^{M}\right)^{1-\sigma_M} P_{Mj}^{\sigma_M-1}E_{Mj}\right] + \frac{E_{Sl}}{w_l} = L_l, \tag{49}$$

These are R non-linear equations in the R wage rates.

D.4 Autarky

It is easy to show that

$$
\begin{aligned}
\frac{L_{Al}}{L_l} &= \frac{E_{Al}}{w_l L_l} = \beta\left(1-\frac{\underline{M}}{\theta_{Ml}}\right)+(1-\beta)\frac{\underline{A}}{\theta_{Al}}, && (50)\\
\frac{L_{Ml}}{L_l} &= \frac{E_{Ml}}{w_l L_l} = \alpha\left(1-\frac{\underline{A}}{\theta_{Al}}\right)+(1-\alpha)\frac{\underline{M}}{\theta_{Ml}}, && (51)\\
\frac{L_{Sl}}{L_l} &= \frac{L_l-(L_{Al}+L_{Ml})}{L_l} = (1-\alpha-\beta)\left(1-\frac{\underline{A}}{\theta_{Al}}-\frac{\underline{M}}{\theta_{Ml}}\right). && (52)
\end{aligned}
$$

D.5 Free Trade

With costless trade, and assuming $\sigma_A=\sigma_M=\sigma$, it is easy to show that

$$
\frac{L_{Ml}/L_{Al}}{L_{Ml'}/L_{Al'}} = \left(\frac{\theta_{Ml}/\theta_{Al}}{\theta_{Ml'}/\theta_{Al'}}\right)^{\sigma-1}. \tag{53}
$$

The relative share of the services sector depends positively on the country's wage (as long as the effect of w_l on P_{Al} and P_{Ml} is assumed negligible):

$$
\frac{L_{Sl}}{L_l} = \frac{(1-\alpha-\beta)\left(w_l-\underline{A}P_{Al}-\underline{M}P_{Ml}\right)}{w_l}. \tag{54}
$$

D.6 Costly Trade

Example 1 with three sectors Consider a three-country world, $R=3$, and a geographic structure such that country 1 takes a "central" position while countries 2 and 3, which are fully symmetric, are in the "periphery": we model this by assuming that country 1 can trade with both 2 and 3 at positive but finite trade costs ($T_{12}=T_{21}=T_{13}=T_{31}=T>1$) and that countries 2 and 3 cannot trade with one another ($T_{23}=T_{32}=\infty$). Trade costs are assumed equal across industries. We simplify further by assuming $\sigma_M=\sigma_A=\sigma$. Finally, assume all parameters are identical across countries (except for the bilateral trade costs) and, in particular, that $\theta_{Aj}=\theta_{Mj}=\theta_{Sj}=L_j=1$ and $\underline{M}=0$. Profiting from the symmetry we have imposed, let us normalize $w_2=w_3=1$.

The results discussed in example 1 above apply here as well.

Example 2 with three sectors Assume all parameters are identical across countries (except for the bilateral trade costs) and, in particular, that $\theta_{Aj}=\theta_{Mj}=\theta_{Sj}=L_j=1$, $\sigma_A=\infty$, and $\sigma_M>1$ but finite. We simplify further by assuming $\underline{A}=\underline{M}=0$. Again, we consider a geographic structure such that country 1 takes a "central" position while countries 2 and 3 are in the "periphery": here we model this by assuming that country 1 can trade freely with both 2 and 3 ($T_{12}=T_{21}=T_{13}=T_{31}=1$) and that countries 2 and 3 cannot trade with one another ($T_{23}=T_{32}=\infty$). Trade costs are equal across sectors here. We take the agricultural good as the numéraire.

The results discussed in example 2 above apply here as well with small variations. First, it is easy to show that $L_{Sj}=1-\alpha-\beta$ for all countries. Second, it is easy to show as well that

$L_{M1} > L_{M2} = L_{M3}$. Third, if parameter values yield $L_{A1} = 0$, then the labor market equilibrium conditions yield $w_1 > w_2 = w_3 = 1$.

D.7 Calibration

Since the third sector is non-tradable and non-differentiated, we do not require new estimates for σ_M, σ_A, δ_{1M}, δ_{1A}, δ_{2M}, and δ_{2A} (note that the expression for manufacturing and agricultural exports in the model is the same as in the two sector version). However, we do require new estimates for α, $\underline{A}$, β, and $\underline{\mathrm{M}}$ since β and $\underline{\mathrm{M}}$ are new parameters and the meaning of α and $\underline{A}$ has changed due to the introduction of the third sector.

To obtain estimates of these new parameters, we follow our earlier approach to use expenditure shares and food consumption for the richest and poorest country in our data, respectively. To see this, note that expenditure shares in the new model are given by:

$$
\begin{aligned}
S_{EMj} &= \frac{E_{Mj}}{w_j L_j} = \alpha - \alpha \underline{A} \frac{P_{Aj}}{w_j} + (1-\alpha)\, \underline{\mathrm{M}} \frac{P_{Mj}}{w_j}, \\
S_{EAj} &= \frac{E_{Aj}}{w_j L_j} = \beta + (1-\beta)\, \underline{A} \frac{P_{Aj}}{w_j} - \beta \underline{\mathrm{M}} \frac{P_{Mj}}{w_j} \\
S_{ESj} &= \frac{E_{Sj}}{w_j L_j} = (1-\alpha-\beta) - (1-\alpha-\beta)\, \underline{A} \frac{P_{Aj}}{w_j} - (1-\alpha-\beta)\, \underline{\mathrm{M}} \frac{P_{Mj}}{w_j}
\end{aligned}
$$

As $w_j \to \infty$, $S_{EMj} \to \alpha$, $S_{EAj} \to \beta$ and $S_{ESj} \to (1-\alpha-\beta)$. Thus, a suitable proxy for α and β are the expenditure shares of the richest country in the data.

Likewise, note that as $w_j \to (\underline{A} P_{Aj} + \underline{\mathrm{M}} P_{Mj})$, consumption per head convergence to the agricultural and manufacturing subsistence levels:

$$
\begin{aligned}
\lim_{w_j \to (\underline{A} P_{Aj} + \underline{\mathrm{M}} P_{Mj})} \frac{E_{Aj}}{L_j P_{Aj}} &= \lim_{w_j \to (\underline{A} P_{Aj} + \underline{\mathrm{M}} P_{Mj})} \left(\underline{A} + \alpha P_{Aj}^{-1} w_j - \alpha \underline{A} - \alpha \underline{\mathrm{M}} P_{Mj} P_{Aj}^{-1} \right) = \underline{A}, \\
\lim_{w_j \to (\underline{A} P_{Aj} + \underline{\mathrm{M}} P_{Mj})} \frac{E_{Mj}}{L_j P_{Mj}} &= \lim_{w_j \to (\underline{A} P_{Aj} + \underline{\mathrm{M}} P_{Mj})} \left(\underline{\mathrm{M}} + \mu P_{Mj}^{-1} w_j - \beta \underline{A} P_{Aj} P_{Mj}^{-1} - \beta \underline{\mathrm{M}} \right) = \underline{\mathrm{M}},
\end{aligned}
$$

Since $w_j = \underline{A} P_{Aj} + \underline{\mathrm{M}} P_{Mj}$ is the income level which guarantees that the substistence level is just attainable, a suitable proxy for $\underline{A}$ and $\underline{\mathrm{M}}$ are real expenditure per worker in the poorest country in our data (measured in \$I).[46]

E Appendix E: Estimating Substitution Elasticities

The demand side structure in Broda and Weinstein is very similar to ours. In particular, they define a composite imported good M_t which aggregates individual goods in a CES fashion:

$$
M_t = \left(\sum_{g \epsilon G} (M_{gt})^{(\gamma_g - 1)/\gamma_g} \right)^{\gamma_g / (\gamma_g - 1)}
$$

[46] Note that since we define "rich" and "poor" as total expenditure per worker (which is consistent with our model), the ranking of countries changes slightly with the introduction of a third sector (services expenditure is now taking into account in the definition of income). That is, the poorest country according to a new definition is now Tanzania, explaining the increase in the estimate for $\underline{A}$ in Table 3.

M_{gt} is the subutility derived from the consumption of imported good g at time t. Note that in our setting, we only have two such goods (the manufacturing and agriculture composite good) and that we assume a Cobb-Douglas rather than a CES aggregator. This does not matter in the following because we are interested in substitution elasticities at the next lower level of aggregation only. Similar to us, Broda and Weinstein assume that M_{gt} aggregates varieties differentiated by country of origin and that, in addition, the aggregator takes the following nonsymmetric CES form:

$$M_{gt} = \left(\sum_{c \epsilon C} d_{gct}^{1/\sigma_g} \left(m_{gct}\right)^{(\sigma_g - 1)/\sigma_g} \right)^{\sigma_g/(\sigma_g - 1)} \tag{55}$$

where σ_g is the elasticity of substitution among varieties of good g and d_{gct}^{1/σ_g} denotes a taste or quality parameter for a variety from country c.[47] Associated with this aggregator is the price index:

$$\Phi_{gt}^{M} = \left(\sum_{c \epsilon C} d_{gct} \left(p_{gct}\right)^{1-\sigma_g} \right)^{1/(1-\sigma_g)}$$

where p_{gct} is the price charged by country c for good g at time t. From (55), we can derive the following import demand function (expressed as import shares and in log-differences):

$$\Delta \ln s_{gct} = \varphi_{gt} - (\sigma_g - 1) \Delta \ln p_{gct} + \varepsilon_{gct}$$

where $\varphi_{gt} = (\sigma_g - 1) \ln \left(\Phi_{gt}^M / \Phi_{gt-1}^M \right)$ and $\varepsilon_{gct} = \Delta \ln d_{gct}$.

Broda and Weinstein also allow for an upward-sloping supply curve of the form:

$$\Delta \ln p_{gct} = \psi_{gt} + \frac{\omega_g}{1+\omega_g} \Delta \ln s_{gct} + \delta_{gct}$$

where $\omega_g \geq 0$ is the inverse supply elasticity, $\psi_{gt} = \frac{\omega_g}{1+\omega_g} \ln E_{gt}$, E_{gt} is total expenditure on good g at time t in the importing country, and $\delta_{gct} = \frac{1}{1+\omega_g} \Delta \ln \upsilon_{gct}$ captures random changes in the technology factor υ_{gct}. Crucially for the identification strategy below, Broda and Weinstein further assume that demand and supply shocks are independent, implying $E\left(\varepsilon_{gct}\delta_{gct}\right) = 0$.

Supply and demand can be rewritten to eliminate the intercepts φ_{gt} and ψ_{gt} by normalizing with respect to a reference country k:[48]

$$\varepsilon_{gct}^{k} = \Delta^k \ln s_{gct} + (\sigma_g - 1) \Delta^k \ln p_{gct}$$

$$\delta_{gct}^{k} = \Delta^k \ln p_{gct} - \frac{\omega_g}{1+\omega_g} \Delta^k \ln s_{gct}$$

where $\Delta^k \ln p_{gct} = \Delta \ln p_{gct} - \Delta^k \ln p_{gkt}$, etc. To take advantage of $E\left(\varepsilon_{gct}\delta_{gct}\right) = 0$, we multiply the two normalized equations and obtain:

[47] Feenstra (1994, p.161) shows that allowing for quality differences is important to address the aggregation problem arising from the fact that we only observe unit values rather than prices in the trade data. One problem resulting from this is that we implicitly ignore changes in the number of varieties supplied from each exporting country (we have assumed this away for simplicity in our model but such changes are likely to be an important phenomenon in the data). However, Feenstra demonstrates that changes in the number of varieties are isomorphic to changes in the quality parameters d_{gct}, and thus captured by the error term ε_{gct} in the regression to be estimated below.

[48] We choose the reference country so that the number of usable observations is maximised (we need share price data for both country c and the reference country k). We use the U.S. and Canada as reference countries for the cross-country sample and the U.S. import sample, respectively.

$$\left(\Delta^k \ln p_{gct}\right)^2 = \theta_1 \left(\Delta^k \ln s_{gct}\right)^2 + \theta_2 \left(\Delta^k \ln p_{gct} \Delta^k \ln s_{gct}\right) + u_{gct} \tag{56}$$

with $\theta_1 = \frac{\omega_g}{(1+\omega_g)(\sigma_g-1)}$ and $\theta_2 = \frac{\omega_g(\sigma_g-2)-1}{(1+\omega_g)(\sigma_g-1)}$. Although $u_{gct} = \varepsilon_{gct}\delta_{gct}$ is correlated with shares and prices, we can obtain consistent estimates of θ_1 and θ_2 by implementing the following between estimator (averaging across periods t):

$$\left(\overline{\Delta^k \ln p_{gc}}\right)^2 = \theta_1 \left(\overline{\Delta^k \ln s_{gc}}\right)^2 + \theta_2 \left(\overline{\Delta^k \ln p_{gc} \Delta^k \ln s_{gc}}\right) + \overline{u}_{gc} \tag{57}$$

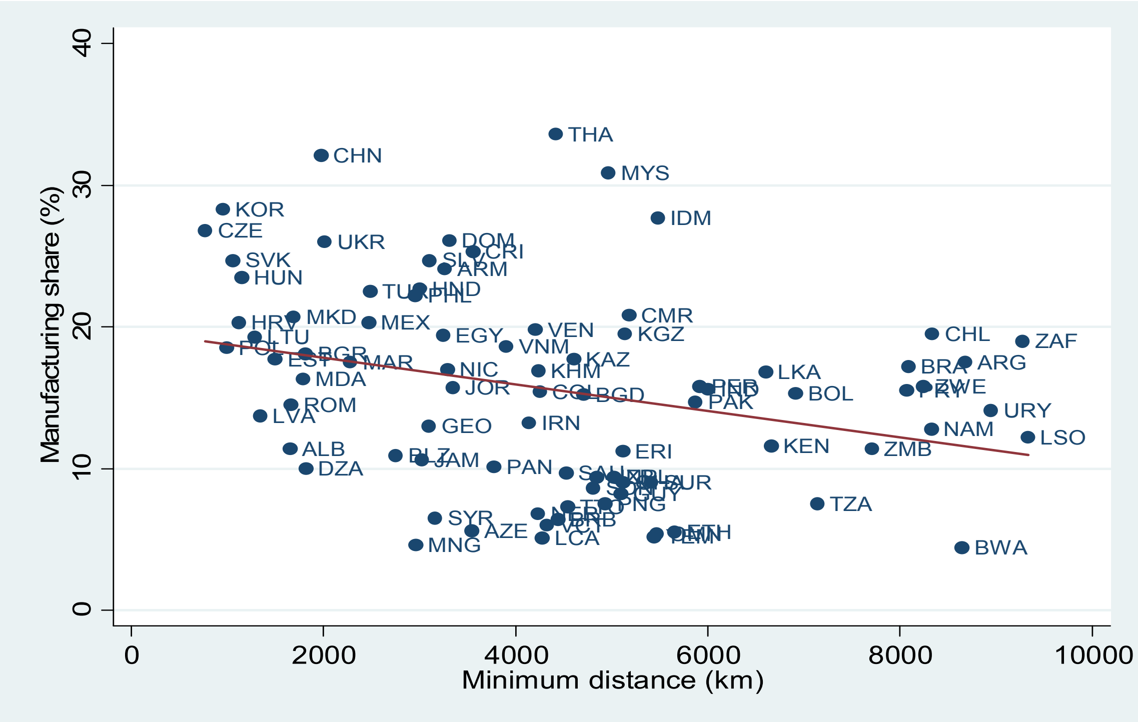
By the assumption of independence of ε_{gct} and δ_{gct}, we know that $E\left(\overline{u}_{gc}\right) = 0$ and thus $plim\left(\overline{u}_{gc}\right) = 0$ as the number of periods T approaches infinity. So the error term in (57) vanishes, solving the problem of correlation with the regressors. We estimate (57) using weighted least squares to obtain estimates for θ_1 and θ_2.[49] Using the definition of θ_1 and θ_2 above, we then solve for ω_g and σ_g.[50]

[49] We follow Broda and Weinstein (2006, pp. 582-584) in adding an additional term inversely related to the quantity of imports from a given country on the right-hand side of (57) and in weighting the data so that the variances are more sensitive to price movements based on large import quantities than small ones. Broda and Weinstein show that this helps addressing problems arising from measurement error due to the use of unit values (rather than actual prices).

[50] If this approach produces economically infeasible estimates (i.e., $\sigma_g \leq 1$ or $\omega_g < 0$), Broda and Weinstein propose to do a grid search over a large set of feasible values. Fortunately, we did not encounter this problem in our estimation.

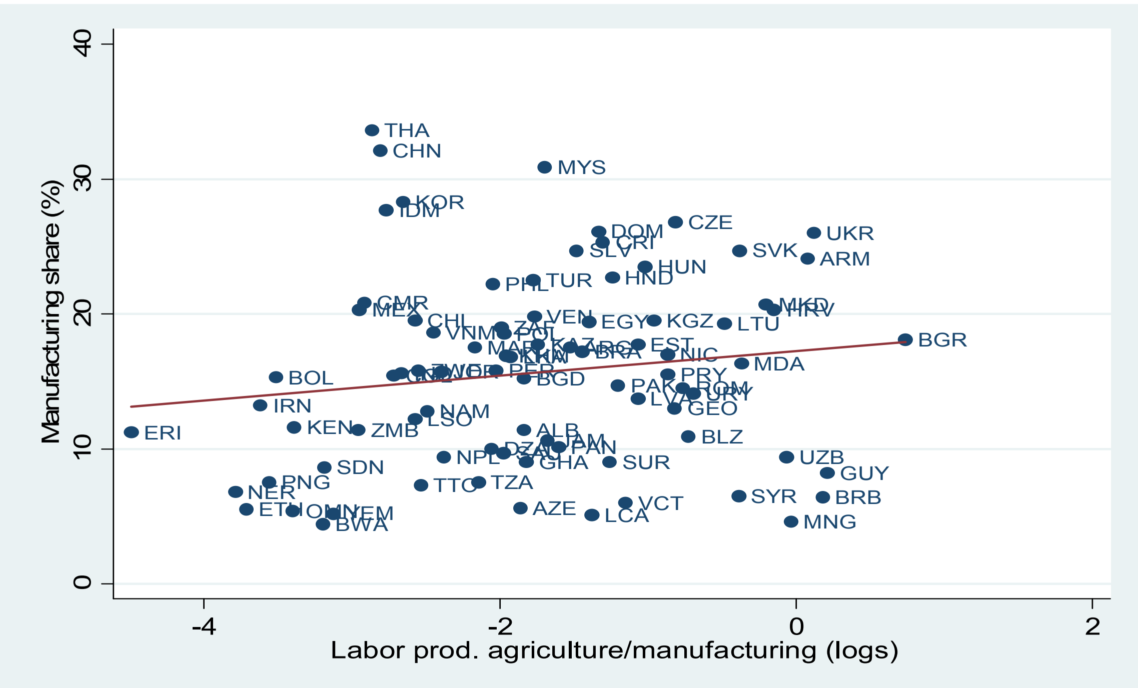
Figures and Tables

Figure 1: GDP manufacturing shares and minimum distance to main markets (2000)



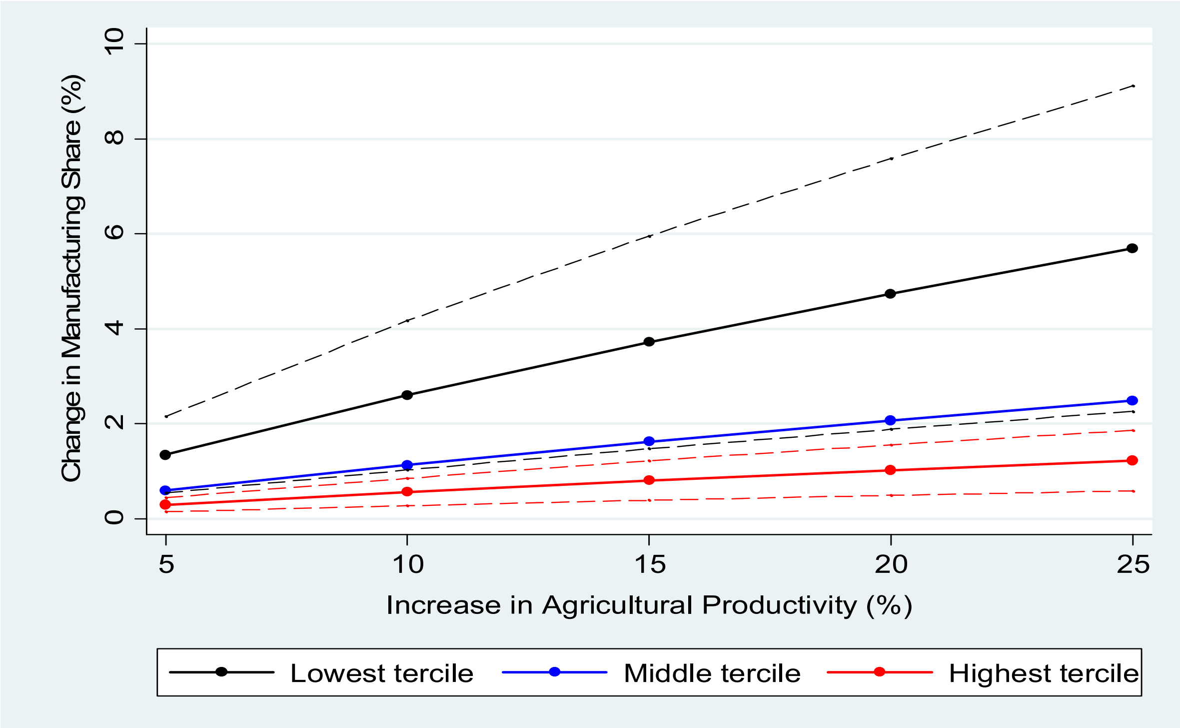
Notes: Figure plots manufacturing shares in GDP (in %) against the minimum distance (in km) to either of the U.S., the European Union (Netherlands), or Japan. All data are for 2000. See Appendix A for data sources and country codes.

Figure 2: GDP manufacturing shares and relative productivities (2000)



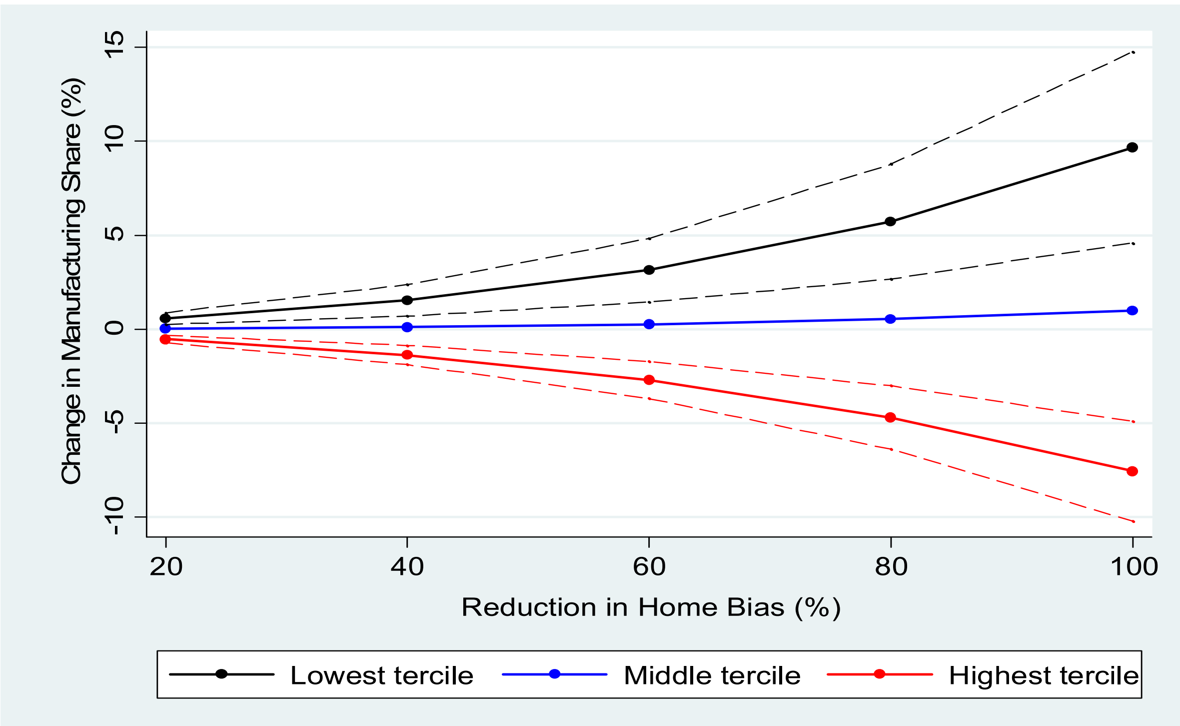
Notes: Figure plots manufacturing shares in GDP (in %) against the ratio of labor productivity in agriculture relative to manufacturing. Labor productivity is measured as value added per worker, adjusted for cross-country price differences using sector-specific PPP exchange rates – see Section 2 and Appendix B for details. All data are for 2000. See Appendix A for data sources and country codes.

Figure 3: Impact of Increases in Agricultural Productivity by Tercile of Centrality



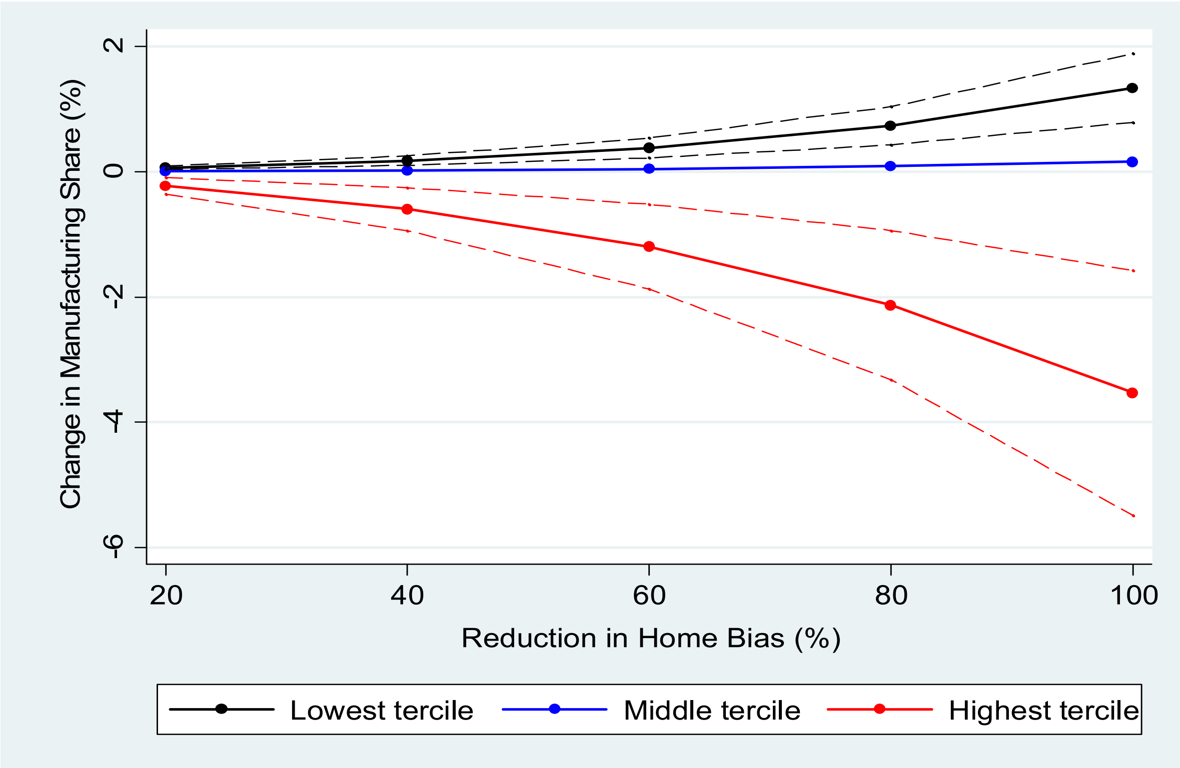
Notes: Figure plots the simulated change in manufacturing GDP shares (in %) due to increases in agricultural productivity, starting from baseline parameter values (see Table 3). Agriculture productivity is increased separately for each country and the manufacturing share change in that country is recorded. Solid lines show averages of the resulting change across countries in the lowest, middle and upper tercile of the centrality measure (see (2) and Section 4 for details). Dotted lines are 95% confidence intervals (omitted for the middle tercile).

Figure 4: Impact of Increases in Openness by Tercile of Comp. Adv. (Global Liberal.)



Notes: Figure plots the simulated change in manufacturing GDP shares (in %) resulting from a global reduction in the home bias parameters δ_{2A} and δ_{2M}, starting from baseline parameter values (see Table 3). Solid lines are averages of the resulting share changes for countries in the lowest, middle and upper tercile of comparative advantage in agriculture, respectively, measured as productivity in agriculture divided by productivity in manufacturing (see Section 4). Dotted lines are 95% confidence intervals (omitted for the middle tercile).

Figure 5: Impact of Increases in Openness by Tercile of Comparative Advantage (Regional Liberalization)



Notes: Figure plots the simulated change in manufacturing shares in GDP (in %) resulting from a regional reduction in the home bias parameters δ_{2A} and δ_{2M}, starting from baseline parameter values (see Table 3). Solid lines are averages of the resulting share changes for countries in the lowest, middle and upper tercile of comparative advantage in agriculture, respectively, measured as productivity in agriculture divided by productivity in manufacturing (see Section 4). Dotted lines are 95% confidence intervals (omitted for the middle tercile).

Table 1: Baseline Empirical Results (Developing Countries Only)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Regressor	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM
log(RELPR)	0.0850				0.129**		0.113**	
	(0.0630)				(0.0585)		(0.0454)	
log(mindist)		-0.332***						
		(0.0754)						
log(CEN)			0.417***	0.434***		0.462***		0.547***
			(0.126)	(0.120)		(0.119)		(0.133)
Fixed Effects	--	--	--	--	SE-Asia Dummy	SE-Asia Dummy	Year	Year
Years	2000	2000	2000	2000	2000	2000	1980, 1990, 2000	1980, 1990, 2000
Observations	83	83	83	80	83	83	256	256
R-squared	0.026	0.118	0.073	0.082	0.242	0.274	0.045	0.093

Notes: Table displays coefficients and robust standard errors (clustered at the country level in columns 7-8) for OLS estimations. The dependent variable is the logistic transformation of a country's share of manufacturing in GDP. RELPR is the quotient of a country's agricultural labor productivity and its labor productivity in manufacturing (labor productivity is defined as value added per worker, adjusted for cross-country price differences using sector-specific PPP exchange rates – see Section 2 and Appendix B for details). Mindist is the minimum distance (in km) of a country to either of Japan, the European Union (Netherlands) or the USA. CEN is a country's centrality measure (defined in Section 2). All regressorsare in logs. Results on the included constant are suppressed. For data sources see Appendix A. *, **, and *** signify statistical significance at the 10%, 5% and 1% levels.

Table 2: Extended Empirical Results (Developing Countries Only)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Regressor	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM
log(RELPR)	0.126*	-0.0307	0.0596	0.116*	0.140***	-0.0348	0.0279	0.0972**			
	(0.0635)	(0.0800)	(0.0746)	(0.0644)	(0.0449)	(0.0540)	(0.0519)	(0.0459)			
log(CEN)	0.344***	0.359**	0.305**	0.255**	0.413***	0.370**	0.281**	0.298**	0.400***	2.588***	2.930***
	(0.129)	(0.148)	(0.121)	(0.122)	(0.129)	(0.142)	(0.113)	(0.114)	(0.137)	(0.881)	(0.980)
log(POP)	0.178***		0.182***	0.195***	0.171***		0.181***	0.187***	0.186***	0.953***	0.602
	(0.0269)		(0.0280)	(0.0289)	(0.0247)		(0.0235)	(0.0233)	(0.0225)	(0.265)	(0.481)
log(AP)		0.0859	0.115*			0.190***	0.220***				
		(0.0703)	(0.0672)			(0.0655)	(0.0557)				
log(PCGDP)				0.135**				0.197***	0.190***	0.289***	0.0854
				(0.0541)				(0.0523)	(0.0469)	(0.101)	(0.170)
Fixed Effects	--	--	--	--	Year	Year	Year	Year	Year	Year, Country	Long First Difference
Years	2000	2000	2000	2000	1980, 1990, 2000	1980, 1990, 2000	1980, 1990, 2000	1980, 1990, 2000	1980-2005	1980-2005	1980-2005
Observations	83	83	83	83	256	256	256	256	2,977	2,977	73
R-squared	0.321	0.088	0.346	0.379	0.305	0.175	0.399	0.398	0.340	0.877	0.145

Notes: Table displays coefficients and robust standard errors (clustered at the country-level in columns 5-11) for OLS estimations. The dependent variable is the logistic transformation of a country's share of manufacturing in GDP. RELPR is the quotient of a country's agricultural labor productivity and its labor productivity in manufacturing (labor productivity is defined as value added per worker, adjusted for cross-country price differences using sector-specific PPP exchange rates – see Section 2 and Appendix B for details). CEN is a country's centrality measure (defined in Section 2). POP is a country's population size, AP its labor productivity in agriculture and PCGDP its per-capita GDP, respectively. All regressors are in logs. Results on the included constant are suppressed. For data sources see Appendix A. *, **, and *** signify statistical significance at the 10%, 5% and 1% levels.

Table 3: Calibrated Parameter Values

Parameter	Value (Baseline)	Value (Robustness)	Outline of Calibration Procedure	Data sources
σ_A	*2.3*	*2.6*	Estimated on cross-country (baseline) and U.S. (robustness) data on import quantities and prices for the year 2000, following Broda and Weinstein (2006) and Broda, Greenfield and Weinstein (2006).	UN-NBER (Feenstra et al., 2005)
σ_M	*2.3*	*2.0*		US-NBER (Feenstra et al., 2002)
θ_{Al}, θ_{Ml}	country- and sector-specific		Labor productivity in manufacturing and agriculture, corrected for cross-country price differences (see Appendix B).	UNIDO, WDI, ICP
$T_{lj}=e^{\delta_1 * d(int)} dist_{lj}^{\delta_2}$	$\delta_{1A}=2.42, \delta_{1M}=1.53,$ $\delta_{2A}=0.74, \delta_{2M}=0.69$	$\delta_{1A}=3.08, \delta_{1M}=1.17,$ $\delta_{2A}=1.05, \delta_{2M}=1.42$	Coefficients on bilateral distance and internal trade flow dummies from gravity equation estimations (baseline: Poisson QML; robustness: OLS), combined with estimates for σ_M and σ_A.	NBER, FAO, CEPII
A, α	$A=170\$I/year$, $\alpha=0.81$		Manufacturing expenditure share of richest country (α) and food expenditure per worker of the poorest country (A) in the sample.	ICP
A, M; α, β	$A=285\$I/year$, $\beta=0.07$ $M=100\$I/year$, $\alpha=0.29$		3-sector model only. Manufacturing and agricultural expenditure share of richest country (α and β) and food and manufacturing expenditure per worker of the poorest country (A and M).	ICP

Notes: Table shows parameter estimates used for the model calibration in Section 4. Also listed are outlines of the calibration procedures and the data sources used (see Section 4 and Appendices B, D and E for details). \$I denotes international dollars.

Table 4: Results for Generated Data (Baseline)

Regressor	(1) ltShareM	(2) ltShareM	(3) ltShareM	(4) ltShareM	(5) ltShareM
log(RELPR)	0.206***		-0.027	-0.023	0.319***
	(0.048)		(0.038)	(0.039)	(0.041)
log(CEN)		0.516***	0.187*	0.181*	0.229**
		(0.156)	(0.097)	(0.093)	(0.088)
log(AP)			0.395***	0.399***	
			(0.047)	(0.051)	
log(POP)				0.007	0.187***
				(0.016)	(0.032)
log(PCGDP)					0.665***
					(0.079)
Observations	79	79	79	79	79
R-squared	0.176	0.145	0.721	0.722	0.801

Notes: Table displays coefficients and robust standard errors for OLS estimations using generated data (see Section 4 for details). The dependent variable is the logistic transformation of a country's share of manufacturing in GDP. RELPR is the quotient of a country's agricultural labor productivity and its labor productivity in manufacturing. CEN is a country's centrality measure (defined in Section 2). POP is a country's population size, AP its labor productivity in agriculture, and PCGDP its per-capita GDP, respectively. All regressorsare in logs. Results on the included constant are suppressed. *, **, and *** signify statistical significance at the 10%, 5% and 1% levels.

Table 5: Results for Generated Data (Developing Countries, Three-Sector Model)

Regressor	(1) ltShareM	(2) ltShareM	(3) ltShareM	(4) ltShareM	(5) ltShareM
log(RELPR)	0.175***		0.002	0.012	0.227***
	(0.047)		(0.032)	(0.034)	(0.046)
log(CEN)		0.501***	0.194**	0.184**	0.197**
		(0.154)	(0.092)	(0.089)	(0.080)
log(AP)			0.277***	0.283***	
			(0.054)	(0.057)	
log(POP)				0.021	0.150***
				(0.019)	(0.038)
log(PCGDP)					0.478***
					(0.091)
Observations	79	79	79	79	79
R-squared	0.164	0.165	0.535	0.541	0.661

Notes: Table displays coefficients and robust standard errors for OLS estimations using generated data (see Section 4 for details). The dependent variable is the logistic transformation of a country's share of manufacturing in GDP. RELPR is the quotient of a country's agricultural labor productivity and its labor productivity in manufacturing. CEN is a country's centrality measure (defined in Section 2). POP is a country's population size, AP its labor productivity in agriculture, and PCGDP its per-capita GDP, respectively. All regressors are in logs. Results on the included constant are suppressed. *, **, and *** signify statistical significance at the 10%, 5% and 1% levels.

Table 6: Results for Generated Data (Robustness Checks for Two-Sector Model)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Regressor	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltshareM	ltShareM	ltShareM
log(RELPR)	0.207***		-0.023	0.231***		0.001	0.206***		-0.027
	(0.047)		(0.038)	(0.048)		(0.036)	(0.048)		(0.038)
log(CEN)		0.502***	0.170*		0.509***	0.139*		0.518***	0.188*
		(0.144)	(0.089)		(0.128)	(0.075)		(0.157)	(0.098)
log(AP)			0.392***			0.397***			0.395***
			(0.046)			(0.048)			(0.047)
Observations	79	79	79	79	79	79	79	79	79
R-squared	0.181	0.149	0.725	0.210	0.180	0.729	0.176	0.145	0.721
σ_A	2.6	2.6	2.6	2.3	2.3	2.3	2.3	2.3	2.3
σ_M	2.0	2.0	2.0	2.3	2.3	2.3	2.3	2.3	2.3
Trade Cost Matrix	Poisson	Poisson	Poisson	OLS	OLS	OLS	Poisson	Poisson	Poisson
Prices Deflators	Consumer	Consumer	Consumer	Consumer	Consumer	Consumer	Producer	Producer	Producer

Notes: Table displays coefficients and robust standard errors for OLS estimations using generated data (see Section 4 for details). The dependent variable is the logistic transformation of a country's share of manufacturing in GDP. RELPR is the quotient of a country's agricultural labor productivity and its labor productivity in manufacturing. CEN is a country's centrality measure (defined in Section 2) and AP its labor productivity in agriculture. All regressors are in logs. Results on the included constant are suppressed. *, **, and *** signify statistical significance at the 10%, 5% and 1% levels.

Table 7: Results for Generated Data (Baseline, Free Trade, Autarky, No Non-Homotheticity)

	Baseline			Free Trade			Autarky			No Non-Homotheticity		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Regressor	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM
log(RELPR)	0.206***		-0.027	-1.300***		-1.300***	0.232***		0.011	-0.020***		-0.022***
	(0.048)		(0.038)	(0.000)		(0.000)	(0.048)		(0.041)	(0.003)		(0.003)
log(CEN)		0.516***	0.187*		-1.960***	0.000		0.038	0.096		-0.021***	0.011*
		(0.156)	(0.097)		(0.282)	(0.000)		(0.110)	(0.085)		(0.007)	(0.006)
log(AP)			0.395***			0.000			0.417***			-0.001
			(0.047)			(0.000)			(0.059)			(0.002)
Observations	79	79	79	79	79	79	79	79	79	79	79	79
R-squared	0.176	0.145	0.721	1.000	0.311	1.000	0.210	0.001	0.727	0.666	0.092	0.685
Interc. (SE) act.	0.000	0.000	0.000	0.431	0.431	0.431	0.040	0.040	0.040	-8.348	-8.348	-8.348
on simul. data	(0.151)	(0.151)	(0.151)	(0.105)***	(0.105)***	(0.105)***	(0.148)	(0.148)	(0.148)	(5.056)*	(5.056)*	(5.056)*
Slope (SE) actual	0.777	0.777	0.777	0.160	0.160	0.160	0.722	0.722	0.722	11.01	11.01	11.01
on simulated data	(0.208)***	(0.208)***	(0.208)***	(0.129)	(0.129)	(0.129)	(0.204)***	(0.204)***	(0.204)***	(6.250)*	(6.250)*	(6.250)*
R^2 actual on simulated data	0.158	0.158	0.158	0.020	0.020	0.020	0.145	0.145	0.145	0.040	0.040	0.040
Corr. (actual, simulated)	0.398	0.398	0.398	0.143	0.143	0.143	0.380	0.380	0.380	0.201	0.201	0.201

Notes: Table displays coefficients and robust standard errors (clustered at the country-level) for OLS estimations using generated data (see Section 4 for details). The dependent variable is the logistic transformation of a country's share of manufacturing in GDP. RELPR is the quotient of a country's agricultural labor productivity and its labor productivity in manufacturing. CEN is a country's centrality measure (defined in Section 2) and AP its labor productivity in agriculture. All regressors are in logs. Results on the included constant are suppressed. *, **, and *** signify statistical significance at the 10%, 5% and 1% levels.

Appendix Tables

Table A.1: Gravity Equation Estimates for 2000 (Poisson and OLS)

	Manufacturing		Agriculture	
	(1)	(2)	(3)	(4)
Regressor	Exports	log(Exports)	Exports	log(Exports)
d_{int}	-1.989***	-1.518***	-3.146***	-4.056***
	(0.119)	(0.368)	(0.164)	(0.578)
log(distance)	-0.906***	-1.842***	-0.962***	-1.355***
	(0.052)	(0.029)	(0.055)	(0.067)
Observations	10,170	10,170	3,145	3,145
R-squared	--	0.836	--	0.716
Estimation method	Poisson	OLS	Poisson	OLS

Notes: Table displays coefficients and robust standard errors (clustered by exporter) for OLS and Poisson QML estimations (see Section 4 for details). The dependent variable is the value of bilateral exports. The regressors are a dummy variable (d_{int}) which takes the value one if a trade flow crosses national borders, and the log of bilateral distance. Also included are a full set of exporter and importer fixed effects. Results on the included constant are suppressed. See Table 2 for data sources. *, **, and *** signify statistical significance at the 10%, 5% and 1% levels.

Table A.2: Results for Actual Data (All Countries)

	(1)	(2)	(3)	(4)	(5)
Regressor	ltShareM	ltShareM	ltShareM	ltShareM	ltShareM
log(RELPR)	0.104*		-0.00138	0.0825	0.112*
	(0.0596)		(0.0754)	(0.0700)	(0.0621)
log(CEN)		0.316***	0.227**	0.215**	0.202**
		(0.0880)	(0.0974)	(0.0903)	(0.0943)
log(AP)			0.0590	0.0662*	
			(0.0403)	(0.0392)	
log(POP)				0.158***	0.160***
				(0.0242)	(0.0245)
log(PCGDP)					0.0641*
					(0.0375)
Observations	112	112	112	112	112
R-squared	0.032	0.069	0.089	0.296	0.300

Notes: Table displays coefficients and robust standard errors for OLS estimations using generated data (see Section 4 for details). The dependent variable is the logistic transformation of a country's share of manufacturing in GDP. RELPR is the quotient of a country's agricultural labor productivity and its labor productivity in manufacturing (labor productivity is defined as value added per worker, adjusted for cross-country price differences – see Section 2 and Appendix B for details). CEN is a country's centrality measure (defined in Section 2). POP is a country's population size, AP its labor productivity in agriculture, and PCGDP its per-capita GDP, respectively. All regressors are in logs. Results on the included constant are suppressed. *, **, and *** signify statistical significance at the 10%, 5% and 1% levels.

Table A.3: Results for Generated Data (All Countries, Three-Sector Model)

Regressor	(1) ltShareM	(2) ltShareM	(3) ltShareM	(4) ltShareM	(5) ltShareM
log(RELPR)	0.184***		0.054*	0.056*	0.164***
	(0.046)		(0.029)	(0.032)	(0.036)
log(CEN)		0.338***	0.089*	0.089*	0.081*
		(0.099)	(0.051)	(0.050)	(0.048)
log(AP)			0.128***	0.128***	
			(0.025)	(0.025)	
log(POP)				0.003	0.083***
				(0.015)	(0.025)
log(PCGDP)					0.284***
					(0.058)
Observations	107	107	107	107	107
R-squared	0.202	0.141	0.411	0.411	0.525

Notes: Table displays coefficients and robust standard errors for OLS estimations using generated data (see Section 4 for details). The dependent variable is the logistic transformation of a country's share of manufacturing in GDP. RELPR is the quotient of a country's agricultural labor productivity and its labor productivity in manufacturing. CEN is a country's centrality measure (defined in Section 2). POP is a country's population size, AP its labor productivity in agriculture, and PCGDP its per-capita GDP, respectively. All regressors are in logs. Results on the included constant are suppressed. *, **, and *** signify statistical significance at the 10%, 5% and 1% levels.