A Simple Theory of Economic Development at the Extensive Industry Margin

Dario Diodato, Ricardo Hausmann, and Ulrich Schetter

CID Faculty Working Paper No. 416
September 2022

Working Papers
Center for International Development
at Harvard University

A Simple Theory of Economic Development at the Extensive Industry Margin*

Dario Diodato
JRC, European Commission
and Harvard Kennedy School
Seville, Spain
dario.diodato@ec.europa.eu

Ricardo Hausmann
Harvard Kennedy School
and Santa Fe Institute
Cambridge, MA
ricardo_hausmann@hks.harvard.edu

Ulrich Schetter
Harvard Kennedy School
Cambridge, MA
ulrich_schetter@hks.harvard.edu

This Version: September 2022

Abstract

We revisit the well-known fact that richer countries tend to produce a larger variety of goods and analyze economic development through (export) diversification. We show that countries are more likely to enter 'nearby' industries, i.e., industries that require fewer new occupations. To rationalize this finding, we develop a small open economy (SOE) model of economic development at the extensive industry margin. In our model, industries differ in their input requirements of non-tradeable occupations or tasks. The SOE grows if profit maximizing firms decide to enter new, more advanced industries, which requires training workers in all occupations that are new to the economy. As a consequence, the SOE is more likely to enter nearby industries in line with our motivating fact. We provide indirect evidence in support of our main mechanism and then discuss implications: We show that there may be multiple equilibria along the development path, with some equilibria leading on a pathway to prosperity while others resulting in an income trap, and discuss implications for industrial policy. We finally show that the rise of China has a non-monotonic effect on the growth prospects of other developing countries, and provide suggestive evidence for this theoretical prediction.

Keywords economic complexity · economic convergence · export diversification · industrial policy · multiple equilibria · poverty trap · product space · structural change

JEL Classification F43 · O11 · O14

*We thank Pol Antràs, Reto Föllmi, Andres Gomez, Will Johnson, Marc Muendler, Frank Neffke, Elias Papaioannou, Romain Wacziarg and seminar participants at the AAG 2019, GCEG 2022, GEOINNO 2022, the Joint Research Centre of the European Commission, the Harvard Growth Lab, and the University of St. Gallen for helpful comments. An earlier version of this paper has been circulated with the title 'Pathways to Prosperity: Economic Development with a Network of Industries'. Ulrich Schetter gratefully acknowledges financial support from the basic research fund of the University of St. Gallen grant 1031513 and from the Swiss National Science Foundation grant IZSEZ0 178724. The views expressed are purely those of the authors and may not in any circumstances be regarded as stating an official position of the European Commission.

1 Introduction

It is well known that industrialized countries produce a larger variety and more sophisticated goods when compared to developing countries. This naturally raises the question how countries can enter new industries and climb the ladder of development. To address this question, we propose a simple theory of economic development at the extensive industry margin and show that its predictions are in line with the data.

Growth at the extensive industry margin is conceptually very different from growth at the intensive margin. Crudely speaking, growth at the intensive margin involves doing more—or better—of the same, while growth at the extensive margin requires doing something different. To analyze what this implies for economic development, we present a tractable small open economy model that is centered on three core presumptions: First, industries differ in their input requirements of technologies, occupations, or tacit know-how, for example, not just at the intensive, but also at the extensive margin. Second, if an input is currently not used domestically, then an economy needs to build up the capability to provide this input first, and building up this capability is costly. Third, if firms invest in building up the capability to provide a certain input, this will eventually spill over to the rest of the economy. We show that these presumptions imply that countries are more likely to diversify into products that require fewer new inputs, and provide indirect evidence in support of our main mechanism. We then argue that this basic observation about economic diversification has potentially profound consequences for development. In particular, in our model, there may be multiple equilibria and strong path dependency along the development path, with certain routes leading to stagnation and others on a pathway to prosperity. As a consequence, there is potentially large scope for industrial policy in our model.

We begin our analysis with presenting novel stylized facts on growth at the extensive industry margin in Section 2. To fix ideas, we focus on inputs from a heterogeneous set of occupations or tasks—a quintessential non-tradeable input—, but our basic logic equally applies to other sorts of non-tradeable inputs. We then document that countries are more likely to enter nearby industries—industries that require less in terms of occupations that are new to a country. This finding is robust to the inclusion of different sets of fixed effects and it is economically significant: Entry in the nearest industries is four times more likely than entry at maximum distance.

To rationalize this pattern, we develop a theory of economic growth at the extensive industry margin in Section 3. We consider a small open economy—the South—that is embedded in a world with many countries at the frontier—the North. There are many industries that differ in their input requirements from a heterogeneous set of occupations. Following Hausmann and Hidalgo (2011) we assume that occupations are non-tradeable inputs, implying that a country can potentially be active in all industries for which the domestic population is capable of performing the set of occupations needed. The North has already developed the capability to perform all occupations, while the South knows how to perform a subset of these occupations only. There are overlapping generations, and pre-existing occupations are freely transferred from one generation to the next, for simplicity. The South grows by building up the capability to perform additional occupations through 'on-the-job' learning, which allows profit-oriented firms to enter new, more sophisticated industries. Such 'on-the-job' learning, however, results in a lower productivity when compared to pre-existing occupations. As a consequence, the productivity of the South is lower in new industries, and more so the more intense an industry is in occupations that are new to the South. In turn, this implies that it is easier for the South to enter industries that are more similar to the South' current activities. Crudely speaking, developing countries cannot jump from producing textiles to producing airplanes, but need to gradually climb the ladder of development by building up the capability to produce in ever more sophisticated industries. Formally, we show that there is a direct mapping between our theory and the stylized facts of Section 2. Our set-up thus provides a simple framework that can rationalize (i) that countries diversify along the development path; and (ii) that they do so by preferentially entering industries which are similar to a country's current activities in terms of their occupational inputs.

To provide additional support for our main meachanism, we consider a variant of our model with a two-stage entry process in Section 4. In this variant, the difficulty of learning new occupations is unknown initially. We show that this uncertainty implies that the probability of the South to survive in a new industry conditional on entry is higher for nearby industries. The underlying mechanism is the same as for the impact of distance on entry: In nearby industries, the overall productivity depends less on the productivity in new occupations. Hence, firms can tolerate a worse random draw for productivity and still find it profitable to stay in the market. We exploit our empirical framework of Section 2 to show that this latter implication is also supported by the data.

Our theory has profound consequences for economic development. We discuss these in Section 5. In our economy, entry in one industry impacts the South's prospect of entering other industries in two ways. First, entry has a positive general equilibrium effect on the wage which lowers profits in other industries. Second, entry trains the domestic population in additional occupations and this facilitates future entry in all industries that make use of these new occupations. The former effect implies that there can be multiple equilibria with respect to which industries the South enters. The latter effect gives rise to strong path dependency, implying that equilibrium selection can have long-lasting consequences: Depending on the structure of the occupational requirements by industry, it is possible that entry in some industries results in stagnation, while entry in others leads the economy on a pathway to prosperity.

A key insight that emerges from these discussions is that the underlying network of occupational inputs introduces an inter-industry externality of entry. As a consequence, there is potentially large scope for industrial policy in our model, which may—or, in case of industry-specific publicly provided inputs, has to—favor entry in some industries over entry in others. We discuss this in Section 5.2.

Lastly, our theory implies that the rise of China has a non-monotonic effect on growth in other developing countries. In Section 5.3, we show that initially, while China is lagging behind, the rise of China has a positive effect on growth in other developing countries as it increases competition in their exporting industries and, hence, causes downward pressure on the wage rate in these countries. When China leapfrogs, however, this has a discrete negative effect on the growth prospects of other developing countries. The reason is that now the industries that these countries may grow into are more competitive. We show that the data is in line with these theoretical predictions.

Related Literature

We start from the observation that—apart from maybe the very top—richer countries tend to be more diversified in terms of their exports, and countries, as they develop, tend to start exporting new, more sophisticated goods (Imbs and Wacziarg, 2003; Cadot et al., 2010; De Benedictis et al., 2009; Parteka and Tamberi, 2013; Cadot et al., 2013; Kehoe and Ruhl, 2013; Brummitt et al., 2020). We then analyze economic development at the extensive industry margin.

The extensive margin features prominently in innovation-based endogenous growth models, including expanding-variety models (Romer, 1987, 1990), quality-ladder models (Grossman and Helpman, 1991c; Aghion and Howitt, 1992), and task-based growth models (Acemoglu and Restrepo, 2018). In these models, goods (or tasks) are symmetric. By contrast, our paper is centered on analyzing *which* industries countries grow into. Moreover, we consider an open economy model with international trade and focus on convergence of developing countries.

We add to the large literatures on economic convergence on the one hand (e.g. Barro et al. 1991; Barro and Sala-I-Martin 1992; Howitt 2000; Hsieh 2002; Aghion et al. 2005; Acemoglu et al. 2006; Rodrik 2012; Gersbach et al. 2013; Peters and Zilibotti 2021), and growth in open economies on the other (e.g. Grossman and Helpman 1991b; Acemoglu 2003; Galor and Mountford 2008; Nunn and Trefler 2010; Chu et al. 2015; Sampson 2016; Arkolakis et al. 2018; Gersbach et al. 2019; Buera and Oberfield 2020; Jäggi et al. 2021) by considering growth through (export) diversification. Our paper is thus closer to Lucas (1993); Sutton and Trefler (2016); Atkin et al. (2021) who analyze export upgrading in open economies.[1] These papers, however, consider frameworks with a one-dimensional ladder of industries, while we consider a network of industries that we can map to the data. This network gives rise to novel inter-industry spillovers.[2]

Hausmann and Klinger (2006); Hidalgo et al. (2007), previously considered a network of industries—the 'product space'—and showed that countries are more likely to enter industries that are similar to their current activities. They 'take an agnostic approach' treating two goods as similar if they 'tend to be produced in tandem' (Hidalgo et al., 2007, p. 484). As opposed to that, we take a more principled approach and consider a network of industries that is rooted in an underlying network of occupational inputs.[3,4] Moreover,

[1]Hausmann and Rodrik (2003) also consider economic development at the extensive industry margin. They suggest that countries have to learn their productivity in new industries first through a process of costly 'self-discovery'. This self-discovery entails a positive learning externality for future entrants who can observe the revealed success of first-movers and then enter industries with high productivity. We consider a myopic entry decision that may be seen as a reduced form capturing such a positive externality of entry *within* industries. Our key point, however, is different: We argue that entry entails an externality *across* industries, as newly learned occupations may be valuable in other industries as well, potentially giving rise to strong path dependency in development. In that regard our set-up is similar to Hausmann and Klinger (2006, 2007) who present a stylized overlapping generations-model, where entry in an industry facilitates entry in related industries by future generations. Our analysis differs along two key dimensions: First, we analyze equilibrium implications. And second, we consider an underlying network of occupations. This not only helps rationalize inter-industry spillovers but it also allows linking our theory to the data.

[2]The network of industries is also an important difference to the 'product cycle' literature that studies how innovative goods are introduced in the North and then diffuse to the South at later stages, e.g. Vernon (1966); Krugman (1979); Grossman and Helpman (1991a); Stokey (1991); Matsuyama (2000); Foellmi et al. (2018).

[3]A large literature uses similar 'agnostic' approaches to analyze patterns of diversification in the data—see e.g. Hidalgo et al. (2018) and Balland et al. (2022) for reviews. This literature robustly documents a strong pattern of related diversification. To our knowledge, our paper is the first to exploit an underlying notion of relatedness instead of following an agnostic approach.

[4]Note that we consider a horizontal network of industries that share inputs and not a network with input-output linkages as e.g. in Ciccone (2002); Jones (2011); McNerney et al. (2022) who show that those linkages can give rise to large multiplier effects and thus help explain cross-country differences in

we analyze economic diversification in the context of a general equilibrium framework and show that there may be multiple equilibria and poverty traps. In our model, catch-up is thus a 'possibility' and not 'a simple consequence of relative backwardness' (Lucas, 1993, p. 269), i.e., our work may help explain 'growth miracles'.

Growth at the extensive industry margin is a form of structural change and, hence, our paper also relates to a large literature on this topic that dates back at least to Kuznets (1957); Chenery (1960). This literature is typically concerned with the gradual shift of employment shares out of agriculture and into manufacturing and, later, services. It explains these changes either with non-homothetic preferences (e.g. Kongsamut et al. 2001; Gollin et al. 2002; Boppart 2014; Alder et al. 2022) or sectoral differences in technological progress (e.g. Baumol 1967; Acemoglu and Guerrieri 2008).[5] We consider export diversification within manufacturing and our work is thus closer to Ngai and Pissarides (2007); Foellmi and Zweimüller (2008); Buera and Kaboski (2012); Duarte and Restuccia (2020); Duernecker et al. (2021); Buera et al. (2022) who analyze structural change at a more disaggregated level. We add to this literature by focusing on the extensive industry margin in a set-up with a network of industries.[6] Importantly, in our case structural change within manufacturing is a source, not a consequence, of 'technological progress', as a greater diversification of its outputs allows an economy to harness gains from a greater specialization of its inputs.

Lastly, our paper is related to the literatures on multiple equilibria in development and on industrial policy. We discuss this in Section 5.

GDP per capita and growth. Hausmann and Hidalgo (2011) and van Dam and Frenken (2020) derive a network of industries from an underlying network of input requirements of non-tradeable 'capabilities', but they consider probabilistic models and generic capabilities. Johnson (2020) also considers a network of industries that is grounded in an underlying network of occupational inputs. He then studies learning-by-doing at the occupation-level and its implications for inter-industry knowledge-spillovers, the evolution of comparative advantage, and aggregate growth. He does, however, not consider growth at the extensive margin, which is our main focus. Boehm et al. (2022) study firm growth at the extensive industry-margin. They measure the similarity of industries based on their intermediate input requirements and use Indian firm-level data to show that firms tend to start making products whose requirements are similar to their current input mix. They do, however, not consider implications in general equilibrium and for aggregate growth, which is our main focus.

[5] Matsuyama (1992); Caselli and Coleman II (2001); Desmet and Rossi-Hansberg (2014); Eckert and Peters (2018); Matsuyama (2019) embed these ideas in open economy models and analyze implications for structural change and spatial equilibrium.

[6] Foellmi and Zweimüller (2008); Buera and Kaboski (2012) also consider the extensive margin but in set-ups with a one-dimensional ladder of goods.

2 Motivating Facts

In this paper, we start from the well known fact that richer countries tend to produce a larger variety of goods—see literature overview—and analyze economic development at the extensive industry margin. In this section, we present additional motivating facts.

Fact 1. Industries differ in their occupational inputs at the extensive margin

In our model, countries enter new and more sophisticated industries by sequentially building up the capability to perform additional occupations, a quintessential non-tradeable input.[7] We therefore begin our analysis with showing that industries differ largely in their occupational inputs, not just at the intensive but also at the extensive margin. To document this, we use Occupational Employment Statistics (OES) data for 2016 from the United States' Bureau of Labor Statistics (BLS). This data provides us with occupational inputs for 290 industries and 820 occupations.[8] Out of these, we consider the 88 mining and manufacturing industries that we map to international trade data—see Facts 2 and 3 below.

The left panel in Figure 1 shows a histogram of the number of occupations by industry. Industries range from employing 22 different occupations at the low-end of the spectrum (*Leather Tanning*) up to employing 207 different occupations (*Navigational Instruments*). Notably, this is about a quarter of all the existing occupations (and about 42% of the 489 occupations employed in our sample of industries).

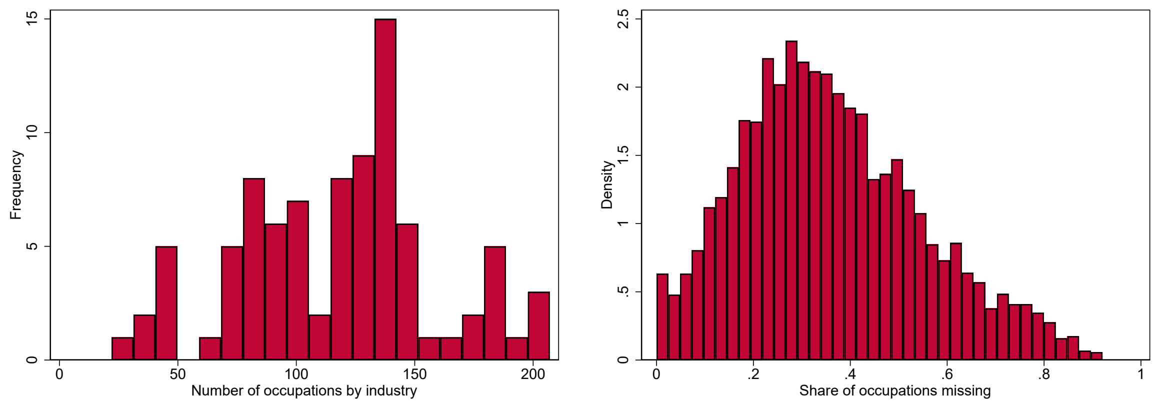
To explore further how diverse (and non-overlapping) input requirements are, we compute for every pair of industries a distance in terms of their occupational input requirements. Specifically, let $\mathcal{T}^i$ denote the set of all occupations employed by industry i, and let $\nu_\tau^i = \mathbb{1}[\tau \in \mathcal{T}^i]$ be an indicator for whether industry i uses occupation τ. With this notation, we compute the following (directional) distance between industries i and j

$$d_{ij} := \frac{\sum_\tau \nu_\tau^j - \sum_\tau (\nu_\tau^i \nu_\tau^j)}{\sum_\tau \nu_\tau^j} \,. \tag{1}$$

[7]Our assumption that the ability to perform occupations needs to be locally available is in line with e.g. Bahar and Rapoport (2018); Hausmann and Neffke (2019); Diodato et al. (2020); Ottinger (2020); Bahar et al. (2022) who show that migrants who bring new skills to an economy are an important driver of economic diversification and structural change. It is also in line with Ellison et al. (2010), who show that occupational similarity is a driver of co-agglomeration of industries across US cities, and with Neffke and Henning (2013) who show that firms are more likely to diversify into industries that make use of their existing human capital.

[8]4-digit North American Industry Classification System (NAICS) and 6-digit Standard Occupational Classification (SOC).

Figure 1: # of occupations by industry (left) and distance between industry pairs (right)



Notes: The left panel shows a histogram of the number of (6-digit SOC) occupations employed by industry (88 4-digit NAICS industries). Y-axis depicts industries' count. The right panel shows the histogram of d_{ij} as defined in Equation (1). Y-axis depicts density of industry pairs. Data source: OES BLS 2016.

$\sum_{\tau} \nu^j_{\tau}$ is the number of occupations required by j and $\sum_{\tau}(\nu^i_{\tau}\nu^j_{\tau})$ is the number of occupations shared by i and j. d_{ij}, in sum, is the distance from industry i to j, in terms of the share of occupations, which j requires but are not used in i.

The right panel in Figure 1 shows the distribution of d_{ij}. It is mildly skewed, with the majority of pairs having between 10% and 60% of occupations missing, and a longer right tail with some having up to 90% missing. This confirms that there is a large degree of variation in the occupational inputs by industry and, hence, suggests large scope for our mechanism of interest.[9] To confirm this, we consider the revealed number of occupations by country next.

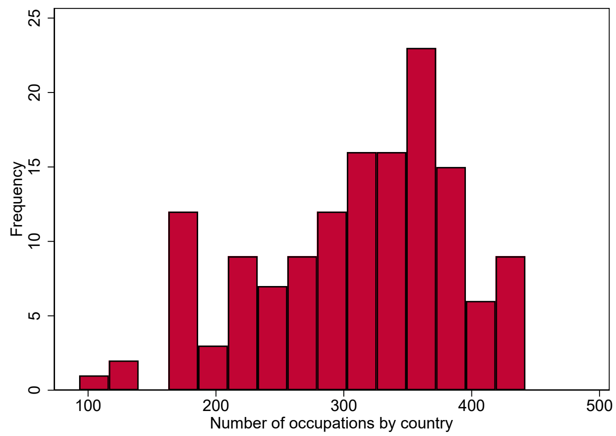
Fact 2. Countries differ in their revealed number of occupations

From the OES, we observe occupational inputs by industry. Following the standard assumption that production technologies are constant across the world—that is, within a given industry, firms use the same set of occupations, irrespective of their location—we can combine this with country-level data on output by industry to back out countries' revealed set of occupations.[10] If, for example, a country is active in two industries, then

[9] The large differences across industries in terms of their occupational inputs also suggests that industries differ largely in their complexity, i.e., there is scope not only for countries to enter new, but also more sophisticated industries. In Appendix C, Tables C.2 and C.3, we show the list of top 10 and bottom 10 industries by number of occupations. This industry ranking is broadly in line with a qualitative notion of high-/low-tech industries.

[10] Ciccone and Papaioannou (2009), for example, have previously used US data to back out industry-specific production functions, with the main focus on human-capital intensity in their case. Likewise, we use US data from OES BLS. We note that our empirical exercise should not find any significant result if the US used an entirely unique technology in production. The fact that the technology in the US provides

Figure 2: Number of occupations by country



Notes: The figure shows a histogram of the revealed number of occupations by country. The graph refers to the period 2012-2016.

we deduce that its population is working on all the occupations that are required by these industries, and treat these occupations as 'present'. In the same spirit, we back out the set of occupations that are present in country c at time t, $\mathcal{T}_{c,t}$.

Production data at the country-industry level is not readily available for a broad set of countries. We therefore use HS6-digit export data from the Atlas of Economic Complexity to proxy for production, and match this to the 88 4-digit NAICS industries in our sample.[11] After cleaning the set of countries to exclude small countries (population < 2 million) and to control for border changes (see Supplementary Material S2.5.2), this allows observing exports by industry for 140 countries. To reduce noise in our data, we pool our data in 5-year windows and then treat a country to be active in an industry if (i) it has a Revealed Comparative Advantage (RCA, Balassa 1965) of at least 1 in (ii) at least 3 years in a 5-year period.[12]

Figure 2 shows a histogram of the revealed number of occupations by country. It clearly shows that a sizable group of countries has many missing occupations. In turn, this allows

a signal on the direction of diversification of other countries—see Fact 3 below—therefore suggests that technology is similar to some extent. Our results are robust to using different vintages for the OES BLS or employment statistics from Mexico—see Supplementary Material S2.1.

[11]Hausmann et al. (2011) adjust the discrepancies between reported importer and exporter values. From this data we retrieve total (global) exports by country for more than 5000 Harmonized System (HS 1992) product codes. Details on the concordance to 4-digit NAICS industries and a robustness check are provided in Supplementary Material S2.1 and S2.5.1.

[12]We introduce this threshold to identify significant and stable exports. A country's Revealed Comparative Advantage in a product is its share in global exports of that product over its share in total world exports (across all products). Using the RCA has the advantage of not being biased against small countries or industries. Robustness checks using other thresholds for the RCA or absolute thresholds for presence are provided in Supplementary Material S2.1.

envisioning development as a process of gradually building up the capability to perform additional occupations and enter new industries. We turn to this next.

Fact 3. Countries are more likely to enter 'nearby' industries

In Hidalgo et al. (2007), it is argued that a country's diversification path is not random but that it is rather determined by overlaps in inputs of non-tradeables capabilities such as skills or know-how. Hidalgo et al. (2007) argue then that the co-occurrence of pairs of products in countries' export baskets can be used as a measure of revealed similarity, and that these similarities can be used to predict what countries will produce next. While this agnostic approach led to the robust empirical regularity that countries diversify into related activities—see Hidalgo et al. (2018) and Balland et al. (2022) for reviews—, the underlying mechanisms remain unclear. Here, we therefore expand on their findings by focusing specifically on occupational inputs. This allows examining countries' export diversification in a more principled way, and later on to build a theoretical model of economic development at the extensive industry margin that exactly maps onto our empirical findings.

To examine countries' entry into new industries, we again pool data in 5-year windows—1992-1996, 1997-2001, 2002-2006, 2007-2011, 2012-2016—, leaving us with $88\times140\times5 = 61{,}600$ observations for global exports by country and industry.[13] We then construct the following entry variable $a^i_{c,t}$ (appearance)

$$a^i_{c,t} = \begin{cases} 1 & \text{if } x^i_{c,t-1} = 0 \text{ and } x^i_{c,t} = 1 \\ 0 & \text{if } x^i_{c,t-1} = 0 \text{ and } x^i_{c,t} = 0 \end{cases}, \tag{2}$$

where $x^i_{c,t}$ is an indicator for whether country c was active in industry i at time t using the threshold based on RCAs from above. Notice that we condition $a^i_{c,t}$ on absence in the previous period, $x^i_{c,t-1} = 0$, so to compare country-industry cells that appeared with those that did not, as we are interested in entry patterns.

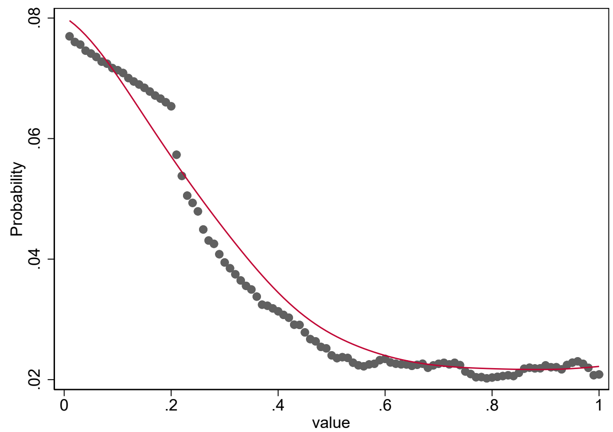
Our main interest is in how entry is related to distance. To that end, we exploit information at both the extensive and the intensive occupational-input margin and measure time-t distance of industry i from country c as

$$\tilde{\mu}^i_{c,t} = \sum_{\tau\in\mathcal{T}\setminus\mathcal{T}_{c,t-1}} \mu^i_\tau. \tag{3}$$

$\mu^i_\tau \geq 0$ is the wage-bill share of occupation τ in industry i, $\mathcal{T}$ is the set of all occupations, and $\mathcal{T}\setminus\mathcal{T}_{c,t-1}$ the set of occupations that were not in use in country c in the previous

[13] See Table C.1 in Appendix C for descriptive statistics.

Figure 3: Probability of industry appearance and distance



Notes: The horizontal axis represents percentiles of $\tilde{\mu}^i_{c,t}$ (p/100). The vertical axis is a moving average of appearances ($a^i_{c,t}$, as in Equation (2)) for an interval ($\pm.2$) around the corresponding x-axis value. The trend line depicts a LOWESS smooth.

period as described above. The use of wage-bill shares is consistent with a Cobb-Douglas production function in our theoretical model. Using employment shares instead yields very similar results—see Table 1 below and Figure S1 in Supplementary Material S2.1.

In Figure 3, we plot the probability of appearance of new industries as the mean $a^i_{c,t}$ for moving intervals of distance $\tilde{\mu}^i_{c,t}$. This figure clearly reveals a sizable negative effect of distance on appearance. At maximum distance—the top-percentile has a distance of 0.7—Figure 3 indicates a probability of appearance of about 2%. At distance zero, when there are no occupations missing in the country, the probability of appearance is 8%, about 4 times larger than at maximum distance, and about 1.6 times larger than the unconditional probability (5%, the mean of $a^i_{c,t}$).

To document that this negative relationship is not driven by country, industry, or time effects, we next run the following regressions

$$a^i_{c,t} = \beta_1 \tilde{\mu}^i_{c,t} + \boldsymbol{\delta}^i_{c,t} + \epsilon^i_{c,t} \ , \tag{4}$$

where $a^i_{c,t}$ and $\tilde{\mu}^i_{c,t}$ are as previously defined, $\boldsymbol{\delta}^i_{c,t}$ is a vector of dummies as specified in the last row of Table 1, and $\epsilon^i_{c,t}$ is an error term. These regressions robustly confirm that entry is inversely related to distance. Notably, the regressions are highly significant in specifications including country×time dummies (countries enter the industries that are nearby), including industry×time dummies (industries appear in nearby countries), and including country×industry dummies (countries enter an industry only once it is close enough). Our theoretical model suggests we use country and industry dummies or country×time and

Table 1: Probability of industry appearance and distance

	(1)	(2)	(3)	(4)	(5)
$\tilde{\mu}^i_{c,t}$ (Cobb-Douglas)	-0.099***	-0.202***	-0.052**	-0.062***	-0.092***
	(0.015)	(0.018)	(0.020)	(0.018)	(0.020)
Adj. R^2	0.04	0.03	0.20	0.05	0.06
N	38586	38586	37979	38586	38586
$\tilde{\mu}^i_{c,t}$ (Leontief)	-0.082***	-0.170***	-0.044**	-0.053***	-0.076***
	(0.013)	(0.015)	(0.018)	(0.015)	(0.016)
Adj. R^2	0.04	0.03	0.20	0.05	0.06
N	38586	38586	37979	38586	38586
	OLS	OLS	OLS	OLS	OLS
Dummies	ct	it	ci	c,i	ct,it

Notes: The dependent variable is appearance of an industry in t ($a^i_{c,t}$), as defined in Equation (2). The independent variable is defined in Equation (3), with the regressions in the top row measuring μ^i_τ using wage-bill shares (Cobb-Douglas) and in the bottom row using employment shares (Leontief). Country-level cluster robust standard errors in parentheses. Significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

industry×times dummies—see Section 3.3. We report on these regressions in Columns (4) and (5). Further robustness checks are provided in Supplementary Material S2.1.[14]

In short, the data suggests that countries climb the ladder of development by preferentially entering nearby industries, viz. industries that require less in terms of new occupations. In what follows, we rationalize these findings in a theoretical model and explore their implications for development.

3 Model

In this section, we present a theory of economic development at the extensive industry margin. We begin with describing the economic environment.

3.1 Economic Environment

We consider a world that is composed of $n+1$ countries, each of measure $\frac{1}{n+1}$: n perfectly symmetric countries in the North, denoted by a subscript $_N$, and one country in the South, denoted by a subscript $_S$.[15] A particularly interesting case is the one where the South is

[14] In these robustness checks, we use different definitions of appearance, different definitions of distance, different sources for the technology matrix (different year, different country), as well as different choices regarding the construction of the dataset. Further details are provided in Supplementary Material S2.1.

[15] We, thus, assume there are only two types of countries, to keep our focus on the catch-up process of the South with respect to the North. In Section 5.3, we relax this assumption and discuss the catch-up

small, i.e., where $n \to \infty$, and for the main part we consider this case. In Supplementary Material S1.2 we show that our main insights are robust to a generic $n \in \mathbb{R}$. Each country is populated by a continuum of measure 1 of overlapping generations that learn when young and work when old. Production combines non-tradeable occupations $\tau \in \mathcal{T}$, where $\mathcal{T}$ denotes the set of all occupations, in a Cobb-Douglas production process. There are I industries that differ in their input requirements of these occupations. In every country and industry there is a (representative) firm equipped with a distinct variety of measure $\frac{1}{1+n}$, analogous to an Armington (1969) model (Anderson, 1979). The North is at the frontier, that is the North is capable of performing all occupations $\tau \in \mathcal{T}$ and it is active in all industries. The South is active in a subset of industries only and grows by learning additional occupations, which allows firms to enter new industries. There is free trade of all goods in every industry.

Households

In each country there is a representative household who has a nested CES utility. The upper-tier utility over industries $i \in \mathcal{I}$, where $\mathcal{I}$ denotes the set of industries, is Cobb-Douglas with industry-shares $\alpha = 1/I$, which we assume to be constant, for simplicity. The lower-tier is CES with elasticity of substitution $\sigma > 1$ between varieties in any given industry. Total (global) expenditure on variety ω in industry i is

$$x^i(\omega)p^i(\omega) = \left[\frac{p^i(\omega)}{P^i}\right]^{1-\sigma} \alpha Y \ , \tag{5}$$

where $x^i(\omega)$ denotes consumption of variety $\omega \in \Omega^i$ in industry i, $p^i(\omega)$ its price, Y aggregate expenditure, and P^i the CES price index for industry i

$$P^i := \left[\sum_{\omega \in \Omega^i} \frac{1}{1+n} p^i(\omega)^{1-\sigma}\right]^{\frac{1}{1-\sigma}} \ . \tag{6}$$

Each household is composed of two generations: young and old. The representative agent in the old generation inelastically supplies L units of labor and spends her income—labor income plus the household share of aggregate profits—on consumption. She supports the young generation in two ways: she shares consumption and leaves the per capita shareholdings of domestic firms as a bequest. The representative agent in the young generation learns and acquires a skill level $\varphi_c \in (0, 1)$ in each occupation that she can observe from the old generation, i.e., in each occupation $\tau \in \mathcal{T}_c$ where $\mathcal{T}_c$ denotes the set

when there are two types of countries in the 'rest-of-the-world', one of which is developing.

of occupations currently in use in country $c \in \{S, N\}$.[16] She can decide to learn entirely new occupations on-the-job when old, but at a discount $\lambda < 1$ on her skill level for these occupations. This discount implies that productivity is initially lower in industries that make use of new occupations and more so, the more intense production is in new occupations. We will detail these implications next.

Production

Consistent with our motivating facts of Section 2, we assume that production combines non-tradeable occupations in a Cobb-Douglas production function

$$y_c^i = \psi^i \prod_{\tau \in \mathcal{T}} (\varphi_{c,\tau} l_{c,\tau}^i)^{\mu_\tau^i} \ , \quad \sum_{\tau \in \mathcal{T}} \mu_\tau^i = 1 \ . \tag{7}$$

y_c^i denotes total output in industry i of country c, $l_{c,\tau}^i$ is the labor input for occupation τ, $\psi^i := \prod_{\tau \in \mathcal{T}} \mu_\tau^{i\ -\mu_\tau^i}$ is a normalizing constant, and

$$\varphi_{c,\tau} := \begin{cases} \varphi_c & \text{if } \tau \in \mathcal{T}_{c,-1} \\ \lambda \varphi_c & \text{otherwise} \end{cases} .$$

Here and below, we use a subscript $_{-1}$ to denote a variable in the previous period, i.e., $\mathcal{T}_{c,-1}$ is the set of occupations that were in use in country c in the previous period. $\mu_\tau^i \geq 0$ denotes the weight of occupation τ in industry i. Importantly, we allow for $\mu_\tau^i = 0$, which implies that occupation τ is not needed for production in industry i. This will play a central role in our analysis. It is in line with the many zeros at the industry-occupation level in the data—see Fact 1 in Section 2.

Rearranging terms yields for the constant productivity of country c in industry i

$$\tilde{\varphi}_c^i := \varphi_c \lambda^{\tilde{\mu}_c^i} \ , \tag{8}$$

where $\tilde{\mu}_c^i := \sum_{\tau \in \mathcal{T} \setminus \mathcal{T}_{c,-1}} \mu_\tau^i$ is our measure of *distance* between country c and industry i: the total importance of occupations that country c needs to newly learn in order to start operating in industry i. This measure of distance is at the heart of our analysis. It exactly maps onto the measure used for Fact 3 from Section 2 as we discuss further below. Observe from Equation (8) that productivity is decreasing in distance. Other than

[16] Essentially, we are assuming that the representative household is perfectly mobile across occupations. Note that we are assuming a long-run perspective here, that is, the young generation learns occupations that it anticipates to work on when old. Hence, we can conceive of each young as learning just one occupation. Then, if the young are a-priori equally skilled in every occupation, foreword-looking behavior on behalf of the young guarantees that the equilibrium wage is the same for every occupation. This is also the case with heterogeneity of households in terms of their (occupation-neutral) skill level.

that, there are no comparative advantages across industries. This allows simplifying the exposition without impacting our main insights.[17]

Firms and Entry

Markets are monopolistically competitive. In every country and industry, there is a (representative) firm equipped with a distinct variety of measure $\frac{1}{n+1}$, i.e., the total measure of varieties is always 1, irrespective of the number of countries in the North, n.[18] Entering an industry involves a fixed cost of entry of f_c^i units of labor. Fixed cost are distributed across industries according to some distribution function $G_f(f)$ with positive and continuous support. In our model this implies that entry into new industries is stochastic. Fixed cost occur in the first period of entering a new industry only. There are no other fixed cost, for simplicity, implying that once a firm starts operating in industry i it will continue doing so forever.[19]

To simplify the exposition, we assume that the old generation does not care about the future well-being of the young generation.[20] The firm in country c and industry $i \notin \mathcal{I}_{c,-1}$ then decides to enter if the per-period profits are larger than the fixed cost of entry

$$\pi_c^{i,v} \geq f_c^i w_c \ , \tag{9}$$

[17]As shown in Section 2, industries differ largely in the number of occupations they use. This suggests that industries differ in their 'complexity', and we might postulate that the South has a comparative advantage in low-complex industries. The data points to a triangular pattern rather than a ladder pattern of international specialization (Hausmann and Hidalgo, 2011; Bustos et al., 2012; Schetter, 2020). Schetter (2020) shows how this can be rationalized in a 'log-supermodular' world once we allow for quality differentiation in an 'O-ring' production process (Kremer, 1993). All our results also apply when considering this O-ring process—see Supplementary Material S1.2 for a discussion. More generally, introducing comparative advantages other than those induced by the learning would simply require conditioning on these comparative advantages. This would complicate matters without adding substance to our main arguments.

[18]In our base case with n large, this directly implies monopolistic competition. More generally, we can think of a continuum of measure $\frac{1}{1+n}$ of potential entrants by country and industry. The only difference to our base case would be that in such case only a subset of firms may find it profitable to enter. Allowing for this possibility would complicate the exposition without adding substance to our main arguments. Moreover, our analysis directly applies to the case with a continuum of potential entrants by country and industry if entry involves aggregate fixed cost that are sufficiently important—see Supplementary Material S1.1. In this latter case, entry may be prone to coordination failure and this may have profound consequences for policy—see Section 5.2 for a discussion.

[19]We make this simplification to avoid the need to keep track of firm exiting which is not material for our main mechanisms of interest. We will relax this assumption in Section 4 below, where we consider a two-stage entry process.

[20]In Supplementary Material S1.2, we discuss a variant with forward-looking entering firms and argue that countries' entry patterns are qualitatively the same. It is often argued that entry into new industries is prone to market failures and has a positive (learning) externality within industries. Our simplifying assumption of myopic entry can also be seen as a reduced form capturing such externalities within industries.

where here and below we use $\pi_c^{i,v}$ to denote the variable profits in country c and industry i, and w_c to denote the equilibrium wage rate in country c, which is the same for all occupations.

Standard steps yield that all active firms choose the well-known constant mark-up over their marginal cost[21]

$$p_c^i = \frac{\sigma}{\sigma - 1} \frac{w_c}{\tilde{\varphi}_c^i} \ . \tag{10}$$

3.2 Equilibrium

In this section, we describe the equilibrium in our economy. We begin with characterizing the equilibrium wage.

CES preferences along with the fact that the North is fully diversified imply that

$$Y = \frac{\sigma}{\sigma - 1} \left[\frac{n}{n+1} L + \frac{w_S}{n+1} \left(L - \sum_{i \in \mathcal{I}_S \setminus \mathcal{I}_{S,-1}} f_S^i \right) \right] \ , \tag{11}$$

where we have chosen the wage in the North to be the numéraire, $w_N = 1$. In words, aggregate expenditure on the consumption aggregator is equal to the constant mark-up over marginal costs times total payouts to production workers. With this we have that the economy is in equilibrium if labor markets clear in all countries given (5), (6), (7), (8), (10), (11), and the entry decisions of the representative firms in the South, (9). Labor markets clear if

$$\sum_{i \in \mathcal{I}_c} L_c^i = L - \sum_{i \in \mathcal{I}_c \setminus \mathcal{I}_{c,-1}} f_c^i$$

for $c \in \{N, S\}$, where L_c^i denotes the demand for production workers in country c and industry i

$$L_c^i := \frac{y_c^i}{\tilde{\varphi}_c^i} \ , \tag{12}$$

and where, recall, $\mathcal{I}_N \setminus \mathcal{I}_{N,-1} = \emptyset$ by assumption.

Let $I_{S,-1}$ denote the number of industries in the South in the previous period, i.e., $I_{S,-1}$ is the number of elements in $\mathcal{I}_{S,-1}$. With this notation, we can characterize the equilibrium wage in the small open economy ($n \to \infty$) as follows:

[21] Strictly speaking, this mark-up applies only to the SOE—our main case of interest. Our main insights also apply to a 'large' South. In such case, firms would still charge the constant mark-up if we assumed a continuum of firms in each country and industry, each equipped with a distinct variety—see Footnote 18.

Proposition 1 (Equilibrium wage)

The equilibrium wage in the South is given by

$$w_S{}^{\sigma} = \frac{L}{L - \sum_{i \in \mathcal{I}_S \setminus \mathcal{I}_{S,-1}} f_S^i} \left[\frac{\varphi_S}{\varphi_N}\right]^{\sigma-1} \alpha \left[I_{S,-1} + \sum_{i \in \mathcal{I}_S \setminus \mathcal{I}_{S,-1}} \lambda^{\tilde{\mu}_S^i(\sigma-1)} \right]. \tag{13}$$

The proof of Proposition 1 is given in Appendix B.1. The wage is lower in the South because (i) it has lower productivity ($\varphi_S < \varphi_N$), (ii) the South is less diversified, i.e., it is active in a subset of industries only ($\mathcal{I}_S \subset \mathcal{I}$), and (iii) some of these industries are new ($\lambda < 1$). As the South enters new industries, its wage rate increases until it reaches

$$w_S = \left(\frac{\varphi_S}{\varphi_N}\right)^{\frac{\sigma-1}{\sigma}}$$

when the South is fully diversified.

With the wage in the South given, the remaining equilibrium values follow. Equations (6), (8), and (10) and the fact that the North is fully diversified imply for the price index in industry i

$$P^i = \begin{cases} \frac{\sigma}{\sigma-1}\left(\frac{1}{n+1}\left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1} w_S{}^{1-\sigma} + \frac{n}{n+1}\left[\varphi_N\right]^{\sigma-1}\right)^{\frac{1}{1-\sigma}} & \text{if } i \in \mathcal{I}_S \\ \frac{\sigma}{\sigma-1}(\varphi_N)^{-1}\left(\frac{n}{n+1}\right)^{\frac{1}{1-\sigma}} & \text{otherwise} \end{cases}.$$

Combining the above with Equations (5), (8), and (10) yields for the world trade share of the South in industry i

$$\frac{y_S^i p_S^i}{\alpha Y} = \begin{cases} \frac{\left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1} w_S{}^{1-\sigma}}{\left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1} w_S{}^{1-\sigma} + n[\varphi_N]^{\sigma-1}} & \text{if } i \in \mathcal{I}_S \\ 0 & \text{otherwise} \end{cases}. \tag{14}$$

3.3 Development at the Extensive Industry Margin

We are now in a position to characterize economic development at the extensive industry margin. In our set-up, the South grows if profit maximizing firms enter new, more advanced industries and the set of industries in the South expands. The South benefits from this diversification as it allows mitigating the decreasing marginal utility in consumption of varieties offered by the South.[22] As a consequence, the real wage in the South increases.

[22] As an alternative, we could consider a case where the South offers many varieties in each industry and where the productivity of the South is heterogeneous across varieties within a given industry as in a multi-sector Eaton and Kortum (2002) model or in a multi-sector Melitz (2003) model. In such case, aggregate productivity increases as the South enters new industries because it allows relocating production from low-productivity varieties in pre-existing industries to high-productivity varieties in new industries.

In short, our model features our key motivating starting point: a positive relation between (export) diversification and income.[23]

With entry into new industries being a key driver of economic development in the South, an ensuing question is which industries the South will enter and when. The remainder of the paper is centered on this question.

With CES preferences variable profits are a constant fraction of revenues. Hence, Equations (11), (14), and the fact that the North is fully diversified imply that a firm in the South finds it profitable to enter industry i whenever

$$\pi_S^{i,v} = \frac{\left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1} {w_S}^{1-\sigma} \alpha \left[nL + w_S \left(L - \sum_{\hat{i} \in \mathcal{I}_S \setminus \mathcal{I}_{S,-1}} f_S^{\hat{i}}\right)\right]}{\left[\left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1} {w_S}^{1-\sigma} + n \left[\varphi_N\right]^{\sigma-1}\right] (\sigma - 1)} \geq f_S^i w_S \ .$$

In the small open economy, this reduces to

$$\frac{\pi_S^{i,v}}{w_S} = \left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1} {w_S}^{-\sigma} \left[\varphi_N\right]^{1-\sigma} \frac{\alpha L}{\sigma - 1} \geq f_S^i \ . \tag{15}$$

Observe from (15) that for any pair of industries that the South might enter, $i_1, i_2 \in \mathcal{I}$: $\tilde{\mu}_S^{i_2} > \tilde{\mu}_S^{i_1}$, variable profits would be higher in industry i_1 than in industry i_2 if wages were kept constant. As we show in Appendix B.2, this is true also when accounting for general equilibrium feedback effects on wages, which in turn immediately implies that the South is more likely to enter nearby industries.

Proposition 2

Consider any pair of industries $i_1, i_2 \in \mathcal{I}$ that the South might enter and suppose that $\tilde{\mu}_S^{i_2} > \tilde{\mu}_S^{i_1}$, i.e., industry i_2 is more distant from the South's pre-existing activities when compared to industry i_1. Then the South is more likely to enter industry i_1 than to enter industry i_2.

Proposition 2 is at the heart of the view on economic development put forth in this paper. It predicts that the South is more likely to enter industries that require less in terms of new occupations. To put it crudely, developing countries cannot jump from producing textiles to producing airplanes, but need to gradually navigate the network of industries.

There is a direct mapping from Proposition 2 to Fact 3 in Section 2. Taking logs on both sides of Condition (15) and rearranging terms yields

$$\underbrace{(\sigma - 1)\ln(\varphi_S) - \sigma \ln(w_S)}_{d_c} + \underbrace{(1-\sigma)\ln(\varphi_N) + \ln\left[(\alpha L)/(\sigma - 1)\right]}_{d^i} + (\sigma - 1)\ln(\lambda)\tilde{\mu}_S^i \geq \ln(f_S^i) \ .$$

[23]To see this, note first that Equation (13) immediately implies that the wage in the South relative to the North increases as $I_{S,-1}$ goes up. And second, that the world production possibilities improve with $I_{S,-1}$. Together, this implies that the real wage in the South must increase with $I_{S,-1}$.

With a uniform distribution of $\ln(f_S^i)$ this corresponds to Regression Model (4) with distance measured as the wage-bill share of new occupations and when controlling for country × time and industry × time fixed effects—the latter to capture time variation in industry size or competitiveness that we abstract from in our analysis, for simplicity. Our theory predicts $\beta_1 := (\sigma - 1)\ln(\lambda) < 0$, in line with our empirical findings.

Proposition 2 follows from our view on economic development as summarized in our key presumptions in the introduction, but it does not hinge on our simplifying assumptions: Proposition 2 is robust to forward-looking behavior on behalf of entering firms, to parts of the rest-of-the-world not being at the frontier, to the South being large, to alternative assumptions regarding fixed costs, a two-stage entry process, and to assuming an O-ring process with an endogenous choice of quality that can rationalize the absence of comparative advantages across industries that differ largely in their complexity. See Supplementary Material S1.2 for further details.

In short, the theoretical framework presented here provides a tractable way of rationalizing Fact 3. While this framework is stylized, it rests on general presumptions that are in line with the data. To substantiate this further, we provide additional suggestive evidence in support of our main mechanism next.

4 Evidence in Support of Main Mechanism

At the heart of our previous arguments is the idea that entry is easier in nearby industries because overall productivity in these industries relies less on the productivity in new occupations. In this section we provide indirect evidence in support of this mechanism. To that end, we exploit a simple extension of our model that leaves our main mechanism intact.

In the baseline version of our model, we assume that λ and, hence, the productivity in new industries, is known to the entering firm. In reality, of course, this may not be, and firms learn their productivity gradually over time. In such case the distance of an industry has testable implications on the probability of continued operation conditional on entry, as we now explain.

To stay as close as possible to our baseline set-up, let us assume that the uncertainty concerns only the productivity in the initial learning period. That is, we assume that the productivity discount on new occupations λ is unknown initially, which allows simplifying the discussions but is not essential for the nature of the argument. In particular, suppose that in a first stage firms decide on paying a given share of the overall fixed cost of entry to learn the country-industry specific discount on new occupations, λ_c^i, drawn from a known distribution $G_\lambda(\lambda)$ with support $[0,1]$.[24] In a second stage, they then decide on whether or not to pay the remainder of the fixed cost to start operating. In such a set-up, there are two opposing forces that govern the impact of distance on the probability of survival conditional on entry: On the one hand, in nearby industries, the productivity of entrants is less elastic with respect to λ, i.e., entrants are more capable of bearing a low λ. Ceteris paribus, this increases the probability of survival conditional on entry. On the other hand, for the same reason profits are larger and, hence, entry is profitable for higher fixed cost f_S^i. Ceteris paribus, this implies a lower probability of survival conditional on entry.[25] In general, which effect dominates depends on the exact circumstances and, in particular, on the distribution of fixed costs, $G_f(\cdot)$. With a uniform distribution of fixed costs, however, the former effect always dominates. We relegate the details to Appendix A and B.4, and summarize the main insight in the following Proposition:

Proposition 3

Consider a two-stage entry process with initially unknown λ and let $f_S^i \sim U[\underline{f}, \overline{f}]$. Then the probability of starting to operate in the second stage conditional on entering in the first stage is strictly decreasing in $\tilde{\mu}_S^i$.

Proposition 3 provides us with an additional testable implication of our main mechanism. It is also centered on the basic observation that a country's productivity in nearby industries is less sensitive to its productivity in new occupations. To take this prediction to the data, we use our theoretically-consistent measure for distance between a country and an industry, and test its ability to predict the following dependent variable

$$y_{c,t}^i = \begin{cases} 1 & x_{c,t-1}^i = 0 \text{ and } x_{c,t}^i = 1 \text{ and } x_{c,t+1}^i = 1 \\ 0 & x_{c,t-1}^i = 0 \text{ and } x_{c,t}^i = 1 \text{ and } x_{c,t+1}^i = 0 \end{cases}, \tag{16}$$

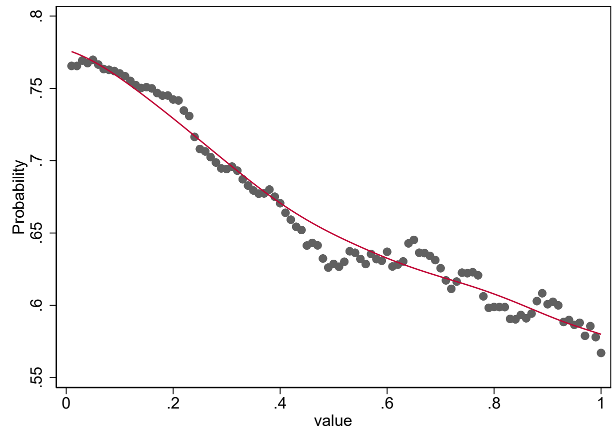
where, again, $x_{c,t}^i$ is an indicator for presence. Notice that we condition $y_{c,t}^i$ on appearance in t. It takes on value of 1 if country c is still active in industry i at time $t+1$ and 0 otherwise.[26]

[24]The restriction that $G_\lambda(\lambda)$ has support $[0,1]$ simplifies the discussion but is not essential. It rules out corner solutions with respect to the probability of starting to operate in the second stage conditional on entering in the first stage (see Appendix A), and it guarantees that this probability is continuous and strictly decreasing in the fixed cost f_S^i. In turn, this allows making *strict* statements about the effect of distance on the probability of starting to operate conditional on entry.

[25]Note that for the same reason the South is more likely to enter nearby industries in this variant of the model as well.

[26]Strictly speaking, in the variant of our model considered here, there is no production in stage 1. Nevertheless, the arguments underlying Proposition 3 are slack, i.e., the result applies to set-ups with

Figure 4: Probability of continued operation and distance



Notes: The horizontal axis represents percentiles of $\tilde{\mu}^i_{c,t}$ (p/100). The vertical axis is a moving average of survival after entry ($y^i_{c,t}$) for an interval ($\pm .2$) around the corresponding x-axis value. The trend line depicts a LOWESS smooth.

Figure 4 shows the raw correlation in the data, analogous to Figure 3. At maximal distance the probability of staying is $\sim 55\%$. This probability increases to 75% at distance zero. Hence, the observed effect of distance is sizeable. Moreover, it is statistically significant and robust to the inclusion of various sets of fixed effects. We show this in Table 2, where we summarize the following sets of regressions

$$y^i_{c,t} = \beta_1 \tilde{\mu}^i_{c,t} + \boldsymbol{\delta}^i_{c,t} + \epsilon^i_{c,t} \ , \tag{17}$$

where $y^i_{c,t}$ and $\tilde{\mu}^i_{c,t}$ are as previously defined, $\boldsymbol{\delta}^i_{c,t}$ is a vector of dummies as specified in the last row of Table 2, and $\epsilon^i_{c,t}$ is an error term. Robustness checks analogous to those for Table 1 are provided in Supplementary Material S2.2.

In summary, in this and the previous section we have presented a simple theory of economic growth at the extensive industry margin that can rationalize our two main motivating facts: (i) Richer countries have more diversified economies and (ii) countries climb the ladder of development by preferentially entering nearby industries. We have provided suggestive evidence in support of our main mechanism. Next, we build on our model and discuss implications of this view on development.

production in the initial ramp-up phase as long as this phase is not too long relative to the overall planning horizon of the entering firm.

Table 2: Probability of continued operation and distance

	(1)	(2)	(3)	(4)	(5)
$\tilde{\mu}^i_{c,t}$ (Cobb-Douglas)	-0.092	-0.957***	-3.143***	-0.448***	-0.479**
	(0.206)	(0.210)	(0.970)	(0.171)	(0.235)
Adj. R^2	0.12	0.04	-0.42	0.13	0.13
N	1486	1529	66	1544	1463
$\tilde{\mu}^i_{c,t}$ (Leontief)	-0.086	-0.809***	-2.936***	-0.382**	-0.406**
	(0.183)	(0.190)	(0.868)	(0.152)	(0.204)
Adj. R^2	0.12	0.04	-0.41	0.13	0.13
N	1486	1529	66	1544	1463
	OLS	OLS	OLS	OLS	OLS
Dummies	ct	it	ci	c,i	ct,it

Notes: The dependent variable is survival after entry of a new industry ($y^i_{c,t}$), as in Equation (16). The independent variable is defined in Equation (3), with the regressions in the top row measuring μ^i_τ using wage-bill shares (Cobb-Douglas) and in the bottom row using employment shares (Leontief). Country-level cluster robust standard errors in parentheses. Significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

5 Implications

5.1 Multiple Equilibria and Path Dependency

In Proposition 2, we have shown that the South is more likely to enter nearby industries. It turns out that the network structure of industries may therefore give rise to multiple equilibria, path dependencies, and income traps at any level of development. In this section, we study these dynamics more carefully.

In our model, entry in one industry has two key effects on firms' potential to enter other industries: First, by introducing new occupations to the economy, entry in new industries facilitates future entry into related industries—industries that make use of these new occupations. This channel captures how the network structure of industries impacts the growth prospects in the South and is our main focus of interest. Second, entry into new industries has general equilibrium effects and, in particular, impacts the wage in the South. This feeds back into the profit potential of firms in all other industries in the current and all future periods. This positive general equilibrium effect on wages implies that entry in different industries are contemporaneously *strategic substitutes*. That is, entry in one industry makes entry in all other industries less attractive, and there may therefore be multiple equilibria. In Appendix B.5, we consider a pair of industries i_1 and i_2 and show that it is possible that entry in either of these two industries is an equilibrium, but not entry in both industries simultaneously.

This multiplicity of equilibria can have profound consequences. In particular, equilibrium selection determines which occupations are available in future and it therefore feeds back into the future prospects of entering new industries. Suppose, for example, that industry i_2 is very isolated in the network of industries. That is, the occupations that need to be newly learned in order to start operating in this industry are exclusive to industry i_2, while industry i_1 is very connected. Then entry into industry i_2 may lead to an income trap, while entry into industry i_1 may lead on a pathway to prosperity. We summarize these insights in the following proposition and provide a formal discussion in Appendix B.5.

Proposition 4

Depending on the structure of occupational inputs by industry,

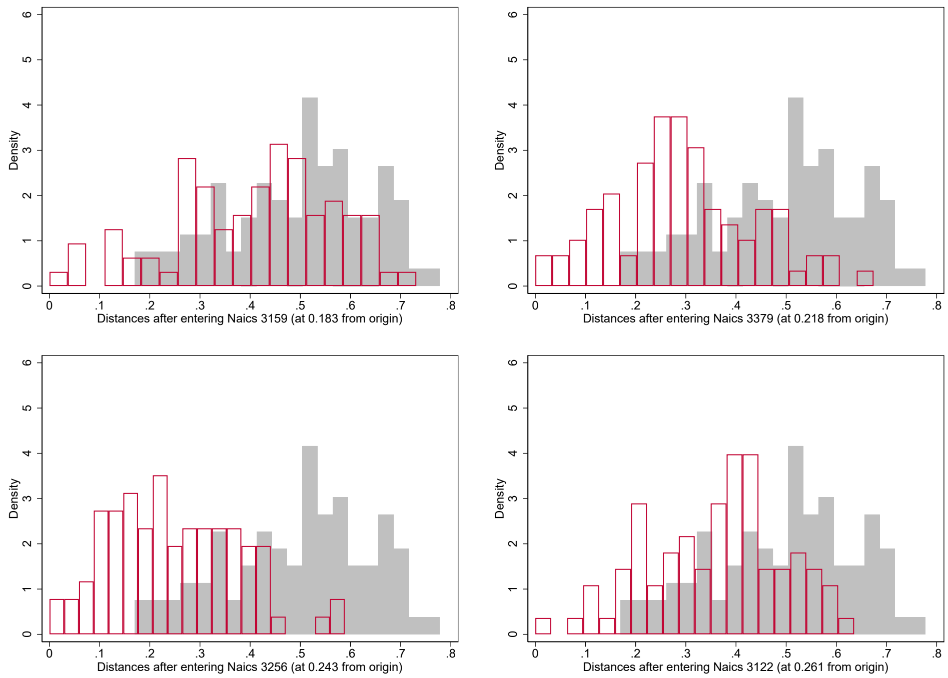
(i) there may be multiple equilibria;

(ii) equilibrium selection today may have long lasting consequences, and there may be income traps at any level of distance from the world technological frontier.

Proposition 4 highlights that with a network of industries there may be multiple equilibria and income traps. Hence, our work also speaks to the tradition of papers arguing that economic development is prone to multiple equilibria—see Matsuyama (1995) for a review of the earlier literature and Allen and Donaldson (2020), Buera et al. (2021), and Choi and Shim (2022) for recent examples. Our main underlying mechanism is, however, different from the canonical one, where multiplicity is rooted in complementarities that arise from industry-wide economies of scale. These complementarities may imply that e.g. it is jointly profitable for many firms to enter an industry, but not so for a single firm. In our case, multiple equilibria arise because entry in different industries are strategic substitutes, and equilibrium selection can have long lasting consequence for development, which are over and above the immediate productivity gains from adopting a 'modern technology' and its spillover effects.

Our coarse data does not allow a rigorous assessment of these insights. We therefore present a descriptive exercise instead that can nevertheless inform the discussions on these issues and point to the potential relevance of the underlying mechanisms. A key element of the arguments underlying Proposition 4 is that industries differ in terms of their connectedness to the rest of the network of industries, particularly the parts not yet occupied by a country. To illustrate this, Figure 5 shows four different possible paths for a country. To keep things simple, we imagine a country c that currently produces in industry 3161 only (Leather Tanning). The solid gray histograms in the background,

Figure 5: Path dependency of entering in new industries

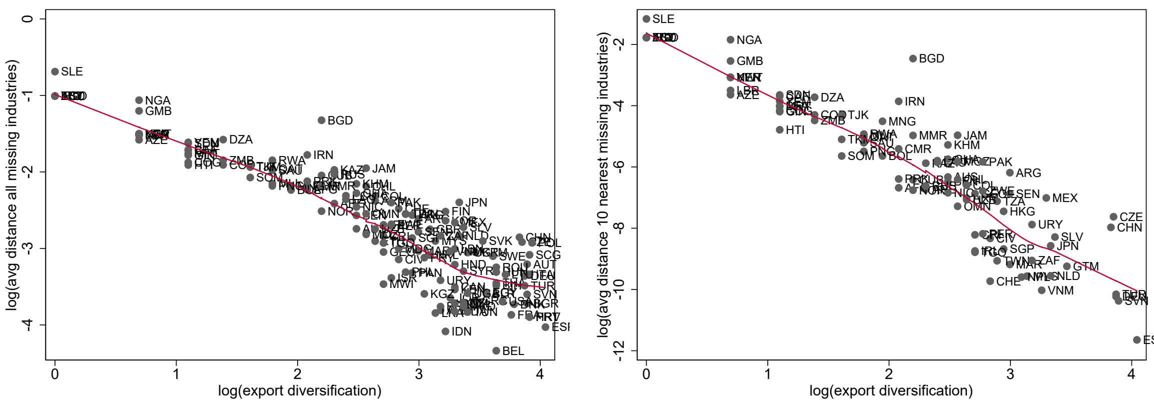


Notes: The histograms show the distribution of distances from a country with a specific set of occupations to all other industries. The solid gray histograms in the background are common across the four panels and represent the distances of all industries from a country that only has industry 3161 (Leather Tanning). The hollow red histograms show how the distances change when the country adds a second industry to its production. The four industries are: Apparel Accessories (3159), Furniture Products (3379), Cleaning Compounds (3256), Tobacco (3122).

which are common across the four panels, depict the distribution of the country's distance $\tilde{\mu}_c^i$ from all other industries, according to OES BLS 2002. Then we ask: how close would that country get to all other industries if we added one more industry? The answer varies significantly, depending on which industry is added. The red hollow histograms in Figure 5 show the distribution of $\tilde{\mu}_c^i$ after the country enters into one of four selected industries. It can be easily observed that the distribution of distances after entering is rather different in the four cases, with two of them (3379 - Furniture Products; 3256 - Cleaning Compounds) experiencing major shifts in the distribution. Entering in industry 3159 (Apparel Accessories), instead, generates only a minor shift, since the occupations used are similar to those of Leather Tanning, which the country already has.[27]

[27] It can, in fact, be noticed that the further away is an industry—distances are indicated in brackets below the histograms and are 0.183, 0.218, 0.243, 0.261, respectively—,the larger the shift in the distribution tends to be. This, however, need not be as, recall, distance is a weighted share of new occupations and because not all skills are equally valuable in other industries. Indeed, if entry is in industry 3122 (Tobacco Manufacturing, the bottom-right panel), the distribution of distances does not shift left as much as in other cases, although the distance is the largest of the four industries considered. It suggests that Tobacco manufacturing uses skills that, even though are new to the country, are not very useful for the production of other industries.

Figure 6: Export diversification and average distance from missing industries



Notes: The horizontal axis is a country's log export diversification using our threshold for a country to be active in an industry as discussed in Section 2 and data from years 2012-2016. The vertical axis is a country's log average distance from its missing industries using wage-bill shares (Cobb-Douglas). The left panel takes the average over all missing industries. The right panel the average over the 10-nearest missing industries. The trend lines depict a LOWESS smooth.

In principle, these arguments are not tied to the stage of economic development of a country, and it is an interesting implication of our theory that income traps can potentially arise at any stage of development. Yet, it is worth noting that a combinatorial argument following Hausmann and Hidalgo (2011) suggests that income traps are more likely to occur at earlier stages of economic development. The reason is that the broader the set of pre-existing occupations, the more options there are for combining any new occupation with pre-existing ones to start operating in a new industry. Similarly, and more to the point for the purpose of our discussions here, this reasoning suggests that countries with a more diversified export basket should be nearer to potential new industries. Figure 6 shows that this is indeed the case in our data. This figure locates countries in a scatterplot with their log-export diversification on the horizontal axis, i.e., the log number of industries that are present in a country according to our criterion from Section 2. The vertical axis represents a country's log-distance from its missing industries based on wage-bill shares. The left panel considers the average distance from all of a country's missing industries. The right panel considers a country's 10-nearest missing industries only. In both figures we observe a strongly negative relationship, i.e., more diversified countries tend to be nearer to potential new industries.[28]

[28]There are fewer countries in the panel on the right because some highly diversified countries have zero distance from their nearest industries. These countries have a log(export diversification) of around 3 or higher, i.e., re-introducing these countries by considering log(average distance $+ \epsilon$) for ϵ small but positive would not change the basic pattern in the figure. In Supplementary Material S2.3, we provide a robustness check where we take averages over a country's missing industries with non-zero distance only. In this case the panel on the right levels off at around log(export diversification) 3, but the figures

5.2 Policy Implications

The previous discussions imply that there is potentially large scope for industrial policy, as we now explain.

The literature on industrial policy—defined as 'imposing tariffs, subsidies, and tax breaks that imply distortions beyond the ones associated with optimal taxes or revenue constraints' (Harrison and Rodríguez-Clare, 2010, p. 4041)—is mostly concerned with (Marshallian) externalities.[29] We add to this literature by focusing on a novel type of inter-industry spillovers.

In our economy with inelastic labor supply and constant mark-ups across industries the equilibrium conditional on entry is efficient (Epifani and Gancia, 2011). Entry, however, entails two externalities. First, entry is myopic and ignores future profits from entering. This simplification can be seen as a reduced form capturing market failures associated with firm entry and positive externalities within industries. It is not essential for our main insights—see Supplementary Material S1.2. Second, however, entry determines the set of occupations that today's young generation can learn and is able to perform without learning costs in future. This, in turn, feeds back into the potential to enter industries in future and constitutes a positive spillover effect that is plausibly an externality of entering and that is the key novelty of our set-up. This externality implies that equilibrium selection can have long lasting consequences and potentially provides significant scope for industrial policy.[30]

Such industrial policy can come in different forms and may also be of the 'soft' type that helps overcome coordination failures and that is advocated e.g. in Harrison and Rodríguez-Clare (2010). To keep things simple, we consider the case of one firm per country-industry pair. As an alternative, we can imagine a continuum of measure $1/(n+1)$ of potential entrants in each country and industry, each equipped with a distinct variety.

look very similar to Figure 6 otherwise, i.e., we still observe a strongly negative relationship between a country's export diversification and its distance from missing industries over the most relevant range of development.

[29]It is well known that in the presence of such externalities, industrial policy can be welfare improving and shift an economy from a 'bad' to a 'good' equilibrium. The empirical evidence is, however mixed—see Harrison and Rodríguez-Clare (2010) and Lane (2020) for surveys of this literature. In a recent contribution, Bartelme et al. (2019), for example, estimate the gains from an optimal system of Pigouvian subsidies to exploit cross-industry heterogeneity in external economies of scale to find gains of 0.61%.

[30]Choi and Levchenko (2021) present a model where industrial policy can have long-run beneficial effects because it stimulates learning-by-doing. In our case, equilibrium selection may have long lasting consequences because there is a network of industries with heterogeneous overlap in terms of input requirements of non-tradeable occupations. We note that this is also different from input-output networks as considered in Blonigen (2016); Liu (2019), for example.

Our analysis from above directly applies to this variant if parts of the fixed cost are industry-wide aggregate fixed cost and if the share of aggregate fixed cost in overall fixed cost is sufficiently large—see Supplementary Material S1.1. The key difference is that in such case a strictly positive mass of firms is needed for it to be profitable to bear the aggregate fixed costs, and there is a need of coordination among potential entrants. Government intervention may help overcoming coordination failures.[31]

Hausmann and Klinger (2006); Hidalgo et al. (2007) previously argued that with a network of industries a country's current pattern of specialization can have profound consequences for development. As opposed to the agnostic approach used in these papers and the related literature, we do not infer industry-similarities based on co-exporting patterns in the data but based on an underlying network of occupational inputs by industry. When it comes to policy, this more principled approach has the advantage of not only allowing to identify target industries for export diversification, but simultaneously informing policy about the missing capabilities—occupational skills in our case—that need to be newly acquired in order to start operating in an industry.

In our model, the South learns new occupations by training its domestic population. As an alternative, the South could acquire the skills needed to enter new industries via migration. Indeed, there is ample evidence suggesting that migration is a key driver of economic diversification and structural change (Bahar and Rapoport, 2018; Hausmann and Neffke, 2019; Diodato et al., 2020; Ottinger, 2020; Bahar et al., 2022). Immigration policy can thus provide a complementary tool to help countries climb the ladder of development.

5.3 Changes in the Competitive Environment—the Ascent of China

Our view on development also provides insights on the impact of the rise of China on the growth prospects of other developing countries.[32] We turn to this issue next.

To shed light on these effects, we consider the case where the rest-of-the-world is composed of two types of countries: the North that is at the frontier and China, which we identify with a subscript $_{CN}$. We assume that a fraction ϕ of countries in the rest-of-the-world is

[31] In addition, production typically relies on publicly provided inputs such as infrastructure or regulation, many of which are industry-specific. If so, governments necessarily have to decide which inputs to provide and how, and this choice feeds into the entry decisions of private firms and may therefore again have long lasting consequences (Hausmann and Rodrik, 2006). The distinct feature of these policy interventions is that they are concerned with supporting the emergence of *new* industries in the South.

[32] See Bloom et al. (2015, 2020); Autor et al. (2020) for recent contributions analyzing the importance of China for growth in industrialized countries.

China. We then ask what happens to the development prospects in the South as China grows and enters new industries. In particular, we assume that China is initially active in a subset of the industries of the South, $\mathcal{I}_{CN} \subseteq \mathcal{I}_{S,-1}$, and then adds industries (and increases its productivity φ_{CN}) until it eventually leapfrogs the South and we have $\mathcal{I}_S \subseteq \mathcal{I}_{CN}$. As we show in Appendix B.6, this has intricate consequences for the development prospects in the South.

Proposition 5

As long as China is lagging behind the South, $\mathcal{I}_{CN} \subseteq \mathcal{I}_{S,-1}$, an increase in either φ_{CN} or $\mathcal{I}_{CN}$ has a positive effect on the growth prospects in the South. When China leapfrogs such that, $\mathcal{I}_S \subseteq \mathcal{I}_{CN}$, this has a discrete negative effect on the growth prospects in the South. Any further increase in φ_{CN} or $\mathcal{I}_{CN}$ has no additional effect on growth in the South.

The basic intuition for the non-monotonic effect of the rise of China on diversification opportunities in the South is simple: Initially, while the South is ahead of China, Chinese growth increases competition in the exporting industries of the South. This puts downward pressure on wages in the South which, in turn, increases the competitiveness of the South in potential new industries. When China leapfrogs, however, these target industries themselves are more competitive and, hence, the profit potential for entering firms in the South is smaller.[33] As China keeps on growing this still increases competition in the South' exporting industries but also increases aggregate demand, and these two effects just offset each other.

To shed light on the empirical relevance of this result and provide additional suggestive evidence in support of our theory, we add two controls to our benchmark regression of appearance in Section 2: 'Competition with China'—the share of goods exported by c that are also exported by China ($\#(\mathcal{I}_{CN} \cap \mathcal{I}_c)/\#(\mathcal{I}_c)$)—and 'Exported by China', which we measure as the share of China in an industry's global exports.[34] Both variables are computed separately for each period, t. Table 3 summarizes the regression results. In line with our theoretical predictions, we find that 'Competition with China' is positively and significantly associated with entry. That is, the presence of China in more of a country's

[33]This is related to the literature analyzing export-biased and import-biased technological change in the tradition of Hicks (1953)—see Grossman and Helpman (1995) for a review. As opposed to this literature, our focus is not on direct welfare implications, but on the impact that technological change in the rest-of-the world (viz. China) has on the diversification prospects of the South.

[34]In these regressions, we control for country, industry, and time dummies (as in our second-strictest specification in Table 1) rather than country×time and industry×time dummies. This is because 'Exported by China' has industry×time variation, while 'Competition with China' has country×time variation.

Table 3: Entry in new industries and the ascent of China

	(1)	(2)	(3)	(4)	(5)	(6)
$\tilde{\mu}^i_{c,t}$ (Cobb-Douglas)	-0.165***	-0.134***	-0.136***			
	(0.015)	(0.017)	(0.017)			
$\tilde{\mu}^i_{c,t}$ (Leontief)				-0.141***	-0.114***	-0.116***
				(0.013)	(0.014)	(0.014)
Exported by China	-0.026***	-0.024***		-0.024***	-0.022**	
	(0.009)	(0.009)		(0.009)	(0.009)	
Competition with China		0.052***	0.053***		0.052***	0.052***
		(0.012)	(0.012)		(0.012)	(0.012)
$Adj.R^2$	0.009	0.011	0.011	0.009	0.011	0.011
Obs.	38381	38381	38381	38381	38381	38381

Notes: The dependent variable is appearance of new industry in t ($a_{ci,t}$), as in Equation (2). The distance variables are defined in Equation (3), with the use of wage-bill share for μ^i_τ denoted by the label 'Cobb-Douglas' and the use of employment share by the label 'Leontief'. 'Exported by China' is the time-t share of China in an industry's global exports. 'Competition with China' is the time-t share of goods exported by c that are also exported by China. Country, industry, and time dummies included in all regressions. Country-level cluster robust standard errors in parentheses. Significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

industries makes it more likely that the country will enter into a new industry. Moreover, the fact that China is a significant exporter in a given industry is negatively associated with entry of other countries.

6 Conclusion

Promoting growth and prosperity in less developed parts of the world remains one of the most pressing challenges of economics and economic policy. Both the cross section of countries today as well as past success-stories of development overwhelmingly suggest that the gradual built-up of the capability to competitively make additional (more complex) products is a key ingredient of economic development. Hence, there is a strong need for a better understanding of the diversification process of countries and for finding ways of promoting it in future.

In this paper, we start from basic facts about economic growth at the extensive industry margin and present a simple theory that can explain them. We provide suggestive evidence in support of our main mechanism. Important lessons emerge from our exercise: Countries are more likely to enter industries that are closer to their current activities in terms of required 'capabilities', and there may be multiple equilibria along the development path, with some leading on a pathway to prosperity and others to stagnation. Hence, there may be large scope for industrial policy, and inevitably so if production requires industry-specific public inputs.

Our work provides a first step to a better understanding of these matters, and many open questions deserve scrutiny in future research. To mention but a few: What is the balance between positive learning-spillovers between industries that use similar ‘capabilities’ and negative pecuniary externalities when competing for the same inputs? Can we identify likely poverty traps in the network of industries and stepping stones on the pathway to prosperity and what is the quantitative importance of our main mechanism? How can we guide policy at a more detailed level regarding which (if any) industries to target and how to support entry? Finding answers to these questions will support better-informed policy making in economic development in future.

Appendix

A Details on Section 4

In this appendix, we provide further details on the extension of our model with a two-stage entry process discussed in Section 4. To capture the main mechanisms of interest and still stay as close as possible to the baseline version of our model, we assume that both stages occur within one period, i.e., one period now consists of two subperiods of equal length, normalized to one, for simplicity. Moreover, we simplify the exposition by considering the case where the South enters in one industry only. Finally, nothing of what follows hinges on the exact share of total fixed cost that needs to be paid in the first stage. We therefore assume that a share $\frac{1}{2}$ of the overall fixed cost are paid in the first and second stage, respectively.

In the first stage, the representative firm in the South weighs the expected gains from entering against the fixed cost of entry, anticipating that it will pay the second half of the fixed cost only if it is profitable to do so. Let $\underline{\lambda}_S^i(f_S^i)$ denote the industry-i threshold level of λ for which it is just profitable to pay the second half of the fixed cost if fixed cost are f_S^i. Clearly, this threshold is implicitly defined by[35]

$$\left[\varphi_S \left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i}\right]^{\sigma-1} \left[w_S(\underline{\lambda}_S^i(f_S^i))\right]^{-\sigma} \left[\varphi_N\right]^{1-\sigma} \frac{\alpha L}{\sigma-1} = \frac{f_S^i}{2} \ , \tag{A.1}$$

where we made explicit that the equilibrium wage w_S depends on the productivity discount in industry i, $w_S(\lambda)$. The representative firm in the South therefore decides to enter industry i in the first stage if

$$\int_{\underline{\lambda}_S^i(f_S^i)}^{1} \frac{\left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1}}{\left[\varphi_N\right]^{\sigma-1}} w_S(\lambda)^{1-\sigma} \frac{\alpha L}{\sigma-1} - \frac{w_S(\lambda) f_S^i}{2} dG_\lambda(\lambda) \geq \frac{w_{S,1} f_S^i}{2} \ , \tag{A.2}$$

where on the right-hand-side of the inequality we use $w_{S,1}$ to denote the wage rate in the first stage, i.e., prior to starting operations in industry i. In words, Condition (A.2) states that the South enters industry i whenever the expected profits net of fixed cost in the second stage are at least as large as the fixed cost in the first stage.

The derivations in the main text imply that for every $\lambda < 1$ and any given f, profits are larger when entering nearby industries. In turn, this immediately implies that the South is more likely to enter nearby industries with unknown λ as well. The two-stage entry

[35] Note that in the second stage of the entry process the fixed cost paid in the first stage are sunk.

process has, however, interesting implications for the probability of starting to operate in stage 2 conditional on entry in stage 1. In particular, consider two industries i_1 and i_2 such that $\tilde{\mu}_S^{i_2} > \tilde{\mu}_S^{i_1}$, i.e., industry i_1 is closer to pre-existing activities in the South when compared to industry i_2. Observe from Equation (13) that $\tilde{\mu}_S^i$ impacts wages only via its effect on $\lambda^{\tilde{\mu}_S^i}$. Equations (13) and (A.1) then imply that for any given $f_S^{i_1} = f_S^{i_2} = f$ it must hold

$$(\underline{\lambda}_S^{i_1}(f))^{\tilde{\mu}_S^{i_1}} = (\underline{\lambda}_S^{i_2}(f))^{\tilde{\mu}_S^{i_2}} \ ,$$

i.e., for a given fixed cost of entry, the firm in industry i_1 is willing to start operating in the second stage at lower levels of λ when compared to industry i_2, $\underline{\lambda}_S^{i_1}(f) < \underline{\lambda}_S^{i_2}(f)$. In turn, this implies that they are more likely to start operating. Intuitively, in industry i_1 a smaller fraction of occupations need to be newly learned and, hence, a productivity discount on new occupations has a smaller impact in industry i_1. For the same reason, however, it is also profitable for the firm in industry i_1 to enter for larger values of the fixed cost when compared to industry i_2. In particular, let $\overline{f}_S^i$ denote the highest fixed cost such that it is just profitable for the firm in industry i to enter in the first stage, i.e., with fixed cost $\overline{f}_S^i$ Condition (A.2) holds with equality. The smaller effect of λ on profits in nearby industries implies that

$$\overline{f}_S^{i_1} > \overline{f}_S^{i_2} \ .$$

Ceteris paribus, higher fixed cost are associated with a higher cutoff $\underline{\lambda}_S^i(\cdot)$ and, hence, a lower probability of starting to operate in the second stage conditional on entering in the first stage. Still, as we show in Appendix B.3, at the respective highest level of fixed cost for which it is just profitable to enter in the first stage, $\overline{f}_S^i$, the probability of starting to operate in the second stage conditional on entering in the first stage is higher in nearby industries:

Lemma 1
$\underline{\lambda}_S^i(\overline{f}_S^i)$ is strictly increasing in $\tilde{\mu}_S^i$.

Lemma 1 typically implies that the probability that the representative firm in the South starts to operate conditional on entering is larger for nearby industries. In particular, this is the case as long as the distribution of fixed cost, $G_f(\cdot)$, is not biased in favor of a lower conditional probability in nearby industries. We demonstrate this in Appendix B.4 where we consider the case of a uniform distribution of f_S^i, $f_S^i \sim U[\underline{f}, \overline{f}]$, but an arbitrary distribution of λ with support $[0, 1]$, and then prove Proposition 3.[36]

[36]For an arbitrary distribution of fixed cost, the probability of operating conditional on entering in the first stage is not necessarily larger for nearby industries because firms in nearby industries are also

B Proofs

B.1 Proof of Proposition 1

Combining (5), (10), and (12) yields for the total demand for labor in the South in industry $i \in \mathcal{I}_S$

$$L_S^i = \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} (P^i)^{\sigma-1} \alpha Y {w_S}^{-\sigma} (\tilde{\varphi}_S^i)^{\sigma-1} \; .$$

Using (6), (8), (10), (11), and the fact that the North is fully diversified yields, after some straightforward rearrangements,

$$L_S^i = \frac{\alpha \left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1} {w_S}^{-\sigma}}{\left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1} {w_S}^{1-\sigma} + n\left[\varphi_N\right]^{\sigma-1}} \left[nL + w_s\left(L - \sum_{\hat{i} \in \mathcal{I}_S \setminus \mathcal{I}_{S,-1}} f_S^{\hat{i}}\right)\right] \; ,$$

where, recall, $\tilde{\mu}_S^i := \sum_{\tau \in \mathcal{T} \setminus \mathcal{T}_{S,-1}} \mu_\tau^i$ denotes the total importance of occupations in industry i that are new to the South. Labor market clearing in the South requires

$$\begin{aligned} \sum_{i \in \mathcal{I}_S} L_S^i &= \left[nL + w_s\left(L - \sum_{i \in \mathcal{I}_S \setminus \mathcal{I}_{S,-1}} f_S^i\right)\right] \sum_{i \in \mathcal{I}_S} \frac{\alpha \left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1} {w_S}^{-\sigma}}{\left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1} {w_S}^{1-\sigma} + n\left[\varphi_N\right]^{\sigma-1}} \\ &= L - \sum_{i \in \mathcal{I}_S \setminus \mathcal{I}_{S,-1}} f_S^i \; . \end{aligned} \tag{B.1}$$

For the small open economy, $\lim_{n \to \infty}$, this simplifies to

$$L\alpha \left[\varphi_N\right]^{1-\sigma} {w_S}^{-\sigma} \sum_{i \in \mathcal{I}_S} \left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1} = L - \sum_{i \in \mathcal{I}_S \setminus \mathcal{I}_{S,-1}} f_S^i \; .$$

Using the fact that $\tilde{\mu}_S^i = 0$ for all $i \in \mathcal{I}_{S,-1}$, this can be rearranged to yield (13).

□

willing to enter in the first stage for higher values of f_S^i. As for any given i, $\underline{\lambda}_S^i(f_S^i)$ is higher for higher f_S^i, higher fixed cost of entry are associated with a lower probability of starting to operate in the second stage conditional on entering in the first stage. In turn, this may imply that the conditional probability of starting to operate is lower in nearby industries. This may be seen most easily when considering the case of a binary distribution of the fixed cost. Suppose, for example, that the fixed cost are 0 with some probability $p > 0$ and $f > 0$ with probability $1-p$. Suppose further that with fixed cost f Condition (A.2) is satisfied for industry i_1 but not for industry i_2 with $\tilde{\mu}_S^{i_2} > \tilde{\mu}_S^{i_1}$. Anticipating this, the representative firm in industry i_2 does not pay the first half of the fixed cost when confronted with f while the representative firm in industry i_1 does. Conditional on entry, the probability of starting to operate is then 1 in industry i_2 but strictly smaller than 1 in industry i_1.

B.2 Proof of Proposition 2

Using Equation (13) in Equation (15) and simplifying terms, we get that it is profitable for the representative firm in industry i to enter if and only if

$$\frac{L - \sum_{\hat{i} \in \mathcal{I}_S \setminus \mathcal{I}_{S,-1}} f_S^{\hat{i}}}{\sigma - 1} \frac{\left[\lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1}}{\sum_{\hat{i} \in \mathcal{I}_S} \left[\lambda^{\tilde{\mu}_S^{\hat{i}}}\right]^{\sigma-1}} \geq f_S^i \ . \tag{B.2}$$

The left-hand-side of Condition (B.2) is decreasing in $\tilde{\mu}_S^i$. Moreover, the left-hand-side of Condition (B.2) is decreasing in f_S^i while the right-hand-side is increasing. Together, this implies that the fixed cost such that it is just profitable for the respective representative firm to enter—i.e., such that Condition (B.2) holds with equality—is lower for industry i_2 than for industry i_1, and the desired result follows from the fact that $G_f(\cdot)$ is increasing.

□

B.3 Proof of Lemma 1

We proceed in three steps. We begin with some preliminary observations. We then show two lemmata. Finally, we use these lemmata to prove the desired result by contradiction.

Dividing both sides of Condition (A.2) by $\frac{f_S^i}{2}$, using Equation (A.1), and simplifying terms we get

$$\int_{\underline{\lambda}_S^i(f_S^i)}^1 \left[\frac{\lambda^{\tilde{\mu}_S^i(\sigma-1)}}{\left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}} \left[\frac{w_S(\underline{\lambda}_S^i(f_S^i))}{w_S(\lambda)}\right]^{\sigma} - 1 \right] w_S(\lambda) dG_\lambda(\lambda) \geq w_{S,1} \ .$$

Using the expression for the equilibrium wage rate, Equation (13), and simplifying terms, we get

$$\int_{\underline{\lambda}_S^i(f_S^i)}^1 \left[\frac{\lambda^{\tilde{\mu}_S^i(\sigma-1)}}{\left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}} \frac{I_{S,-1} + \left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}}{I_{S,-1} + \lambda^{\tilde{\mu}_S^i(\sigma-1)}} - 1 \right] \cdot \left[I_{S,-1} + \lambda^{\tilde{\mu}_S^i(\sigma-1)}\right]^{\frac{1}{\sigma}} dG_\lambda(\lambda) \tag{B.3}$$
$$\geq \left[I_{S,-1}\right]^{\frac{1}{\sigma}} \ ,$$

where here and below we use $I_{S,-1}$ to denote the number of elements in set $\mathcal{I}_{S,-1}$, and where we made use of the fact that—in the current period—the South is entering only industry i.

Note that the right-hand-side of Condition (B.3) is independent of the industry that the South is entering. We therefore focus on the left-hand-side of Condition (B.3), LHS (B.3).

In the following two lemmata, we characterize how LHS (B.3) depends on $\tilde{\mu}_S^i$ and $\underline{\lambda}_S^i(f_S^i)$, respectively.

Lemma 2

The left-hand-side of Condition (B.3) *is strictly decreasing in* $\underline{\lambda}_S^i(f_S^i)$.

Proof Using Leibniz' integral rule, we get

$$\begin{aligned}\frac{dLHS(\text{B.3})}{d\underline{\lambda}_S^i(f_S^i)} =& -\left[\frac{\left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}}{\left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}} \frac{I_{S,-1}+\left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}}{I_{S,-1}+\left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}} - 1\right] \\ & \cdot \left[I_{S,-1}+\left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}\right]^{\frac{1}{\sigma}} g_\lambda(\underline{\lambda}_S^i(f_S^i)) \\ & + \int_{\underline{\lambda}_S^i(f_S^i)}^1 \frac{d}{d\underline{\lambda}_S^i(f_S^i)} \left[\lambda^{\tilde{\mu}_S^i(\sigma-1)} \frac{I_{S,-1}+\left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}}{\left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}\left[I_{S,-1}+\lambda^{\tilde{\mu}_S^i(\sigma-1)}\right]} - 1\right] \\ & \cdot \left[I_{S,-1}+\lambda^{\tilde{\mu}_S^i(\sigma-1)}\right]^{\frac{1}{\sigma}} dG_\lambda(\lambda) \\ & < 0\ .\end{aligned}$$

The inequality follows from noting that the first summand is equal to zero and that

$$\frac{I_{S,-1}+\left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}}{\left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}\left[I_{S,-1}+\lambda^{\tilde{\mu}_S^i(\sigma-1)}\right]}$$

is decreasing in $\underline{\lambda}_S^i(f_S^i)$.

□

Lemma 3

For a given $\underline{\lambda}_S^i(f_S^i)$*, the left-hand-side of Condition* (B.3) *is strictly increasing in* $\tilde{\mu}_S^i$.

Proof Differentiating with respect to $\tilde{\mu}_S^i$ and rearranging terms, we get

$$\begin{aligned}\frac{\partial LHS\ (\text{B.3})}{\partial\tilde{\mu}_S^i} =& \int_{\underline{\lambda}_S^i(f_S^i)}^1 (1-\sigma)\left[I_{S,-1}+\lambda^{\tilde{\mu}_S^i(\sigma-1)}\right]^{\frac{1-\sigma}{\sigma}} \ln(\lambda)\lambda^{\tilde{\mu}_S^i(\sigma-1)} \\ & \cdot \left\{\frac{1}{\sigma} - \frac{\ln(\underline{\lambda}_S^i(f_S^i))}{\ln(\lambda)} + A(\lambda)B(\lambda)\frac{\sigma-1}{\sigma}\right. \\ & \left. + \frac{I_{S,-1}+\left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}}{\left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}}\left[\frac{\ln(\underline{\lambda}_S^i(f_S^i))}{\ln(\lambda)} - 1\right]\right\} dG_\lambda(\lambda)\ ,\end{aligned} \tag{B.4}$$

where we have taken the partial derivative $\frac{\partial LHS(\text{B.3})}{\partial\tilde{\mu}_S^i}$, i.e., holding constant $\underline{\lambda}_S^i(f_S^i)$, and where we introduced the following notation to simplify the exposition

$$A(\lambda) := \left(\frac{\lambda}{\underline{\lambda}_S^i(f_S^i)}\right)^{\tilde{\mu}_S^i(\sigma-1)} \qquad B(\lambda) := \frac{I_{S,-1}+\left[\underline{\lambda}_S^i(f_S^i)\right]^{\tilde{\mu}_S^i(\sigma-1)}}{I_{S,-1}+\lambda^{\tilde{\mu}_S^i(\sigma-1)}}\ .$$

Now, suppose that $I_{S,-1} = 0$. Then $A(\lambda)B(\lambda) = 1$ and it is a matter of simple algebra to verify that the term in curly brackets in Equation (B.4) is equal to zero for every λ. But then the fact that $I_{S,-1} > 0$ implies that the term in curly brackets in Equation (B.4) must be strictly positive for every $\lambda > \underline{\lambda}_S^i(f_S^i)$. This follows from noting first that $B(\lambda)$ is increasing in $I_{S,-1}$ for every $\lambda > \underline{\lambda}_S^i(f_S^i)$ because $\sigma > 1$ and $\underline{\lambda}_S^i(f_S^i) < \lambda \leq 1$, i.e., with $I_{S,-1} > 0$ we must have $A(\lambda)B(\lambda) > 1$ for every $\lambda > \underline{\lambda}_S^i(f_S^i)$. Second, that $\ln(\underline{\lambda}_S^i(f_S^i))/\ln(\lambda) > 1$ for every $\lambda > \lambda_S^i(f_S^i)$. In turn this implies that the last summand in curly brackets is increasing in $I_{S,-1}$, while the first two summands are unaffected by $I_{S,-1}$. Together, this implies that with $I_{S,-1} > 0$, the term in curly brackets is strictly positive for all $\lambda > \lambda_S^i(f_S^i)$. The inequality then follows from noting that the term outside the curly brackets on the right-hand-side of Equation B.4 is strictly positive for all $\lambda < 1$. We conclude that

$$\frac{\partial LHS \text{ (B.3)}}{\partial \tilde{\mu}_S^i} > 0 \ .$$

□

To show Lemma 1, we now proceed by contradiction. Consider a pair of industries i_1 and i_2 such that $\tilde{\mu}_S^{i_2} > \tilde{\mu}_S^{i_1}$, and suppose that $\underline{\lambda}_S^{i_2}(\overline{f}_S^{i_2}) \leq \underline{\lambda}_S^{i_1}(\overline{f}_S^{i_1})$. The definitions of $\underline{\lambda}_S^i(\cdot)$ and of $\overline{f}_S^i$ imply that Condition (B.3) must hold with equality when evaluated at $\underline{\lambda}_S^i(\overline{f}_S^i)$. Let $H^i(\underline{\lambda})$ denote the left-hand-side of Condition (B.3) when evaluated for industry i with distance $\tilde{\mu}_S^i$ and for the case of $\underline{\lambda}_S^i(f_S^i) = \underline{\lambda}$. If $\underline{\lambda}_S^{i_2}(\overline{f}_S^{i_2}) = \underline{\lambda}_S^{i_1}(\overline{f}_S^{i_1})$ then

$$H^{i_2}(\underline{\lambda}_S^{i_2}(\overline{f}_S^{i_2})) > H^{i_1}(\underline{\lambda}_S^{i_1}(\overline{f}_S^{i_1})) \tag{B.5}$$

by Lemma 3. If $\underline{\lambda}_S^{i_2}(\overline{f}_S^{i_2}) < \underline{\lambda}_S^{i_1}(\overline{f}_S^{i_1})$ then

$$H^{i_2}(\underline{\lambda}_S^{i_2}(\overline{f}_S^{i_2})) > H^{i_2}(\underline{\lambda}_S^{i_1}(\overline{f}_S^{i_1})) > H^{i_1}(\underline{\lambda}_S^{i_1}(\overline{f}_S^{i_1})) \ , \tag{B.6}$$

where the first inequality follows from Lemma 2 and the second inequality follows from Lemma 3. Conditions (B.5) and (B.6) are a contradiction to Condition (B.3) holding with equality for both industries, i_1 and i_2, when evaluated at $\underline{\lambda}_S^{i_1}(\overline{f}_S^{i_1})$ and $\underline{\lambda}_S^{i_2}(\overline{f}_S^{i_2})$, respectively. The contradiction establishes the desired result.

□

B.4 Proof of Proposition 3

We proceed in two steps: We first show a preliminary result and then use this to prove Proposition 3.

Let ϵ_S^i denote the semi-elasticity of the ratio of variable profits in industry i over the wage rate in the South with respect to the productivity-discount on new occupations, λ_S^i,

$$\epsilon_S^{i,\pi} := \frac{d \ln\left(\frac{\pi_S^{i,v}(\lambda_S^i)}{w_S(\lambda_S^i)}\right)}{d\lambda_S^i} ,$$

where, again, we have emphasized the dependence of $\pi_S^{i,v}$ and of w_S on λ_S^i. As we show in the following lemma, ϵ_S^i is increasing in $\tilde{\mu}_S^i$.

Lemma 4
$\frac{d\epsilon_S^{i,\pi}}{d\tilde{\mu}_S^i} > 0$

Proof Using Equation (13) in Condition (15), taking logs, and differentiating the left-hand-side with respect to λ_S^i yields, after some straightforward rearrangements,

$$\epsilon_S^{i,\pi} := \frac{d \ln\left(\frac{\pi_S^{i,v}(\lambda_S^i)}{w_S(\lambda_S^i)}\right)}{d\lambda_S^i} = \frac{\tilde{\mu}_S^i(\sigma-1)}{\lambda_S^i}\left[1 - \underbrace{\frac{(\lambda_S^i)^{\tilde{\mu}_S^i(\sigma-1)}}{I_{S,-1} + (\lambda_S^i)^{\tilde{\mu}_S^i(\sigma-1)}}}_{:=D(\lambda_S^i,\tilde{\mu}_S^i)}\right] .$$

Clearly, $0 < D(\lambda_S^i, \tilde{\mu}_S^i) < 1$ and, hence, this semi-elasticity is positive. Differentiating with respect to $\tilde{\mu}_S^i$ and rearranging terms yields

$$\frac{d\epsilon_S^{i,\pi}}{d\tilde{\mu}_S^i} = \frac{\sigma-1}{\lambda_S^i}\left\{1 - D(\lambda_S^i,\tilde{\mu}_S^i) - D(\lambda_S^i,\tilde{\mu}_S^i)\tilde{\mu}_S^i(\sigma-1)\ln(\lambda_S^i)\left[1 - D(\lambda_S^i,\tilde{\mu}_S^i)\right]\right\} > 0 ,$$

where the inequality follows from the fact that $0 < D(\lambda_S^i, \tilde{\mu}_S^i) < 1$ and, hence, the term in curled brackets is positive.

$\square$

In words, Lemma 4 implies that for a given λ_S^i the proportional change of $\frac{\pi_S^{i,v}(\lambda_S^i)}{w_S(\lambda_S^i)}$ in response to a marginal change of λ_S^i is larger for higher values of $\tilde{\mu}_S^i$.

We now prove the statement in Proposition 3. The expected probability of starting to operate in industry i in the second stage conditional on entering in the first stage is

$$\Pr[operate|entry; i] = \int_{\underline{f}}^{\overline{f}_S^i} \left[1 - G_\lambda(\underline{\lambda}_S^i(f))\right] \frac{1}{\overline{f}_S^i - \underline{f}} df .$$

Consider industry i_2. By Lemma 1 we have $\underline{\lambda}_S^{i_2}(\overline{f}_S^{i_2}) > \underline{\lambda}_S^{i_1}(\overline{f}_S^{i_1})$ and, hence, there exists a

$\hat{f} < \overline{f}_S^{i_2}$ such that $\underline{\lambda}_S^{i_2}(\hat{f}) = \underline{\lambda}_S^{i_1}(\overline{f}_S^{i_1})$. $\Pr[operate|entry; i_2]$ can therefore be rewritten as

$$\begin{aligned}\Pr[operate|entry; i_2] =& \frac{\overline{f}_S^{i_2} - \hat{f}}{\overline{f}_S^{i_2} - \underline{f}} \int_{\hat{f}}^{\overline{f}_S^{i_2}} \left[1 - G_\lambda(\underline{\lambda}_S^{i_2}(f))\right] \frac{1}{\overline{f}_S^{i_2} - \hat{f}} df \\ &+ \frac{\hat{f} - \underline{f}}{\overline{f}_S^{i_2} - \underline{f}} \int_{\underline{f}}^{\hat{f}} \left[1 - G_\lambda(\underline{\lambda}_S^{i_2}(f))\right] \frac{1}{\hat{f} - \underline{f}} df \ . \end{aligned} \tag{B.7}$$

Now, $G_\lambda(\underline{\lambda}_S^{i_2}(\hat{f})) = G_\lambda(\underline{\lambda}_S^{i_1}(\overline{f}_S^{i_1}))$ by the definition of $\hat{f}$. Moreover, Lemma 4 along with the definition of $\underline{\lambda}_S^i$ (see Equation (A.1)) imply that $\underline{\lambda}_S^{i_2}(\gamma\hat{f}) > \underline{\lambda}_S^{i_1}(\gamma\overline{f}_S^{i_1})$ and, hence, $G_\lambda(\underline{\lambda}_S^{i_2}(\gamma\hat{f})) > G_\lambda(\underline{\lambda}_S^{i_1}(\gamma\overline{f}_S^{i_1}))$ for every $0 < \gamma < 1$. It follows that

$$\begin{aligned}\int_{\underline{f}}^{\hat{f}} \left[1 - G_\lambda(\underline{\lambda}_S^{i_2}(f))\right] \frac{1}{\hat{f} - \underline{f}} df &< \int_{\underline{f}\frac{\overline{f}_S^{i_1}}{\hat{f}}}^{\overline{f}_S^{i_1}} \left[1 - G_\lambda(\underline{\lambda}_S^{i_1}(f))\right] \frac{1}{\overline{f}_S^{i_1} - \underline{f}\frac{\overline{f}_S^{i_1}}{\hat{f}}} df \\ &< \int_{\underline{f}}^{\overline{f}_S^{i_1}} \left[1 - G_\lambda(\underline{\lambda}_S^{i_1}(f))\right] \frac{1}{\overline{f}_S^{i_1} - \underline{f}} df \ , \end{aligned} \tag{B.8}$$

where the first inequality follows from applying a simple change of variables—$\gamma = f/\hat{f}$ on the left-hand-side and $\gamma = f/\overline{f}_S^{i_1}$ on the right-hand-side—and from then using the above. The second inequality follows from the fact that $\overline{f}_S^{i_1} > \hat{f}$ and therefore $\frac{\overline{f}_S^{i_1}}{\hat{f}}\underline{f} > \underline{f}$ in combination with the fact that $\underline{\lambda}_S^{i_1}(f)$ and therefore $G_\lambda(\underline{\lambda}_S^{i_1}(f))$ is increasing in f. Finally, the fact that $\underline{\lambda}_S^{i}(f)$ is increasing in f also implies that

$$\int_{\hat{f}}^{\overline{f}_S^{i_2}} \left[1 - G_\lambda(\underline{\lambda}_S^{i_2}(f))\right] \frac{1}{\overline{f}_S^{i_2} - \hat{f}} df < \int_{\underline{f}}^{\hat{f}} \left[1 - G_\lambda(\underline{\lambda}_S^{i_2}(f))\right] \frac{1}{\hat{f} - \underline{f}} df \ . \tag{B.9}$$

Conditions (B.7) to (B.9) imply that

$$\Pr[operate|entry; i_1] > \Pr[operate|entry; i_2] \ ,$$

which proves the desired result.

□

B.5 Proof of Proposition 4

We show the proposition by providing an example for each statement. We begin with some preliminary derivations.

From Condition (B.2) we know that it is profitable to enter industry i if and only if

$$\frac{L - \sum_{\hat{i} \in \mathcal{I}_S \setminus \mathcal{I}_{S,-1}} f_S^{\hat{i}}}{\sigma - 1} \frac{\left[\lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1}}{I_{S,-1} + \sum_{\hat{i} \in \mathcal{I}_S \setminus \mathcal{I}_{S,-1}} \left[\lambda^{\tilde{\mu}_S^{\hat{i}}}\right]^{\sigma-1}} \geq f_S^i \ . \tag{B.10}$$

Now, consider the set of industries $\mathcal{I}\backslash\mathcal{I}_{S,-1}$ that the South might potentially enter and any pair of industries $i_1, i_2 \in \mathcal{I}\backslash\mathcal{I}_{S,-1}$. Suppose that for $\hat{i} \in \mathcal{I}\backslash\mathcal{I}_{S,-1}$ it holds

$$\frac{L - f_S^{\hat{i}}}{\sigma - 1} \frac{[\lambda]^{\tilde{\mu}_S^{\hat{i}}(\sigma-1)}}{I_{S,-1} + [\lambda]^{\tilde{\mu}_S^{\hat{i}}(\sigma-1)}} \begin{cases} \geq f_S^{\hat{i}} & \text{if } \hat{i} \in \{i_1, i_2\} \quad \text{(B.11a)} \\ < f_S^{\hat{i}} & \text{otherwise ,} \quad \text{(B.11b)} \end{cases}$$

and that for $\hat{i} \in \{i_1, i_2\}$ it holds

$$\frac{L - f_S^{i_1} - f_S^{i_2}}{\sigma - 1} \frac{[\lambda]^{\tilde{\mu}_S^{\hat{i}}(\sigma-1)}}{I_{S,-1} + \left[\lambda^{\tilde{\mu}_S^{i_1}}\right]^{\sigma-1} + \left[\lambda^{\tilde{\mu}_S^{i_2}}\right]^{\sigma-1}} < f_S^{\hat{i}} \ . \tag{B.12}$$

Then, entering either industry i_1 or i_2 are both equilibria which proves part (i).[37]

Suppose further that any occupation that is newly learned when entering industry i_2 is used exlusively in that industry and let the South enter this industry in the current period, i.e., let $\mathcal{I}_S = \mathcal{I}_{S,-1} \cup \{i_2\}$. Then $\tilde{\mu}_{S,+1}^{\hat{i}} = \tilde{\mu}_S^{\hat{i}}$ for any industry $\hat{i} \neq i_2 \in \mathcal{I}\backslash\mathcal{I}_{S,-1}$, i.e., the distance for these industries is the same in the next period and clearly Condition (B.11b) is still satisfied for all industries $\hat{i} \neq i_1, i_2 \in \mathcal{I}\backslash\mathcal{I}_{S,-1}$. Moreover, if

$$\frac{L - f_S^{i_1}}{\sigma - 1} \frac{\left[\lambda^{\tilde{\mu}_{S,+1}^{i_1}}\right]^{\sigma-1}}{I_S + \left[\lambda^{\tilde{\mu}_{S,+1}^{i_1}}\right]^{\sigma-1}} \ , < f_S^{i_1} \tag{B.13}$$

it is no longer profitable to enter industry i_1 and there will be no further growth, i.e., the South is in an income trap.[38]

Finally, suppose the South enters industry i_1 in the current period, i.e., let $\mathcal{I}_S = \mathcal{I}_{S,-1} \cup \{i_1\}$

---

[37] Note that we can always find parameter values such that Conditions (B.11) and (B.12) are both satisfied. This is because for any $f_S^{i_1}, f_S^{i_2} > 0$ it holds

$$L - f_S^{i_1} - f_S^{i_2} < L - f_S^{i_1}$$

and

$$I_{S,-1} + \left[\lambda^{\tilde{\mu}_S^{i_1}}\right]^{\sigma-1} + \left[\lambda^{\tilde{\mu}_S^{i_2}}\right]^{\sigma-1} > I_{S,-1} + \left[\lambda^{\tilde{\mu}_S^{i_1}}\right]^{\sigma-1} \ .$$

implying that the set

$$\mathcal{F} = \left(\frac{L - f_S^{i_1} - f_S^{i_2}}{\sigma - 1} \frac{\left[\lambda^{\tilde{\mu}_S^{i}}\right]^{\sigma-1}}{I_{S,-1} + \left[\lambda^{\tilde{\mu}_S^{i_1}}\right]^{\sigma-1} + \left[\lambda^{\tilde{\mu}_S^{i_2}}\right]^{\sigma-1}}, \frac{L - f_S^{i_1}}{\sigma - 1} \frac{\left[\lambda^{\tilde{\mu}_S^{i}}\right]^{\sigma-1}}{I_{S,-1} + \left[\lambda^{\tilde{\mu}_S^{i_1}}\right]^{\sigma-1}} \right]$$

is non-empty, and analogous for i_2. Hence, Conditions (B.11) and (B.12) are both satisfied if, for example, $\tilde{\mu}_S^{i_1} = \tilde{\mu}_S^{i_2}$ are chosen such that (B.11) holds with equality for the industry with higher fixed-cost of entry.

[38] Note that Conditions (B.11a) and (B.13) are, for example, both satisfied in the case considered in Footnote 37 if industry i_1 is the industry with higher fixed cost.

and suppose that there exists some industry $i_3 \in \mathcal{I} \backslash \mathcal{I}_S$ such that

$$\frac{L - f_S^{i_3}}{\sigma - 1} \frac{\left[\lambda^{\tilde{\mu}_{S,+1}^{i_3}}\right]^{\sigma-1}}{I_S + \left[\lambda^{\tilde{\mu}_{S,+1}^{i_3}}\right]^{\sigma-1}} \geq f_S^{i_3} \ . \tag{B.14}$$

Then, the South will enter some additional industries in the next period when entering industry i_1 in the current period, which proves part (ii).[39]

□

B.6 Proof of Proposition 5

We begin with some preliminary observations and then consider the four cases according to Table B.1 separately. To simplify the exposition, we will ignore any productivity discounts in new industries as well as the fixed cost of market entry in China.[40]

Table B.1: Different cases for growth of China

	$\varphi_{CN} \uparrow$	$\mathcal{I}_{CN,-1} \subset \mathcal{I}_{CN}$
$\mathcal{I}_{CN} \subseteq \mathcal{I}_{S,-1}$	(i)	(ii)
$\mathcal{I}_S \subseteq \mathcal{I}_{CN}$	(iii)	(iv)

With China accounting for a fraction ϕ of the world, the wage in China is the unique solution to (see Equation (B.1))[41]

$$\phi = \left(\phi + \frac{1-\phi}{w_{CN}}\right) \sum_{i \in \mathcal{I}_{CN}} \alpha \frac{{\varphi_{CN}}^{\sigma-1} {w_{CN}}^{1-\sigma} \phi}{{\varphi_{CN}}^{\sigma-1} {w_{CN}}^{1-\sigma} \phi + (1-\phi) \left[\varphi_N\right]^{\sigma-1}} \ . \tag{B.16}$$

[39]Note that we can always find parameter values such that such an industry exists. In particular, whenever

$$\frac{\left[\lambda^{\tilde{\mu}_{S,+1}^{i_3}}\right]^{\sigma-1}}{I_S + \left[\lambda^{\tilde{\mu}_{S,+1}^{i_3}}\right]^{\sigma-1}} > \frac{\left[\lambda^{\tilde{\mu}_S^{i_3}}\right]^{\sigma-1}}{I_{S,-1} + \left[\lambda^{\tilde{\mu}_S^{i_3}}\right]^{\sigma-1}} \tag{B.15}$$

we can find an $f_S^{i_3}$ such that (B.11b) and (B.14) are both satisfied. Consider, for example, a case where $\mu_S^{i_3} = 1$ and $\mu_{S,+1}^{i_3} = 0$, i.e., industry i_3 makes only use of occupations that are newly learned when entering industry i_1. Then (B.15) reduces to

$$\frac{1}{I_S + 1} > \frac{\lambda^{\sigma-1}}{I_{S,-1} + \lambda^{\sigma-1}} \ ,$$

and for every $\lambda < 1$ we can find an M such that when the number of elements in $\mathcal{I}_{S,-1}$ is larger than or equal to M (B.15) is satisfied.

[40]At the expense of additional notational complexity, such effects can easily be incorporated.

[41]The fact that Equation (B.16) has a unique solution follows from noting that the RHS of (B.16) is continuous and strictly monotonously decreasing in w_{CN} and satisfies $\lim_{w_{CN} \to 0} RHS$ (B.16) $= \infty$ and $\lim_{w_{CN} \to \infty} RHS$ (B.16) $= 0$.

The wage in the South is given by

$$
\begin{aligned}
{w_S}^{\sigma} = (\phi w_{CN} + 1 - \phi)\frac{\alpha L}{L - \sum_{i \in \mathcal{I}_S \backslash \mathcal{I}_{S,-1}} f_S^i} \Bigg[\sum_{i \in \mathcal{I}_S \cap \mathcal{I}_{CN}} & \frac{\left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1}}{{\varphi_{CN}}^{\sigma-1}{w_{CN}}^{1-\sigma}\phi + (1-\phi)\left[\varphi_N\right]^{\sigma-1}} \\
& + \sum_{i \in \mathcal{I}_S \backslash \mathcal{I}_{CN}} \frac{\left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1}}{(1-\phi)\left[\varphi_N\right]^{\sigma-1}} \Bigg] \ ,
\end{aligned} \tag{B.17}
$$

and the variable profits of firms in the South in industry i are (see Equation (15))

$$
\begin{aligned}
\pi_S^{i,v} = & \frac{\alpha L}{\sigma - 1}(\phi w_{CN} + 1 - \phi)\left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1} {w_S}^{1-\sigma} \\
& \cdot \begin{cases} \left[(1-\phi)\left[\varphi_N\right]^{\sigma-1}\right]^{-1} & \text{if } i \notin \mathcal{I}_{CN} \\ \left[{\varphi_{CN}}^{\sigma-1}{w_{CN}}^{1-\sigma}\phi + (1-\phi)\left[\varphi_N\right]^{\sigma-1}\right]^{-1} & \text{otherwise} \end{cases} .
\end{aligned} \tag{B.18}
$$

As before, firms in the South find it profitable to enter industry i if and only if $\frac{\pi_S^{i,v}}{w_S} \geq f_S^i$.

(i) Using Equation (B.17) in Equation (B.18) along with the fact that $\mathcal{I}_{CN} \subseteq \mathcal{I}_{S,-1}$ yields

$$
\begin{aligned}
\frac{\pi_S^{i,v}}{w_S} = & \frac{L - \sum_{\hat{i} \in \mathcal{I}_S \backslash \mathcal{I}_{S,-1}} f_S^{\hat{i}}}{\sigma - 1}\left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1}\left[(1-\phi)\left[\varphi_N\right]^{\sigma-1}\right]^{-1} \\
& \cdot \left[\sum_{\hat{i} \in \mathcal{I}_S \cap \mathcal{I}_{CN}} \frac{\left[\varphi_S \lambda^{\tilde{\mu}_S^{\hat{i}}}\right]^{\sigma-1}}{{\varphi_{CN}}^{\sigma-1}{w_{CN}}^{1-\sigma}\phi + (1-\phi)\left[\varphi_N\right]^{\sigma-1}} + \sum_{\hat{i} \in \mathcal{I}_S \backslash \mathcal{I}_{CN}} \frac{\left[\varphi_S \lambda^{\tilde{\mu}_S^{\hat{i}}}\right]^{\sigma-1}}{(1-\phi)\left[\varphi_N\right]^{\sigma-1}}\right]^{-1} .
\end{aligned} \tag{B.19}
$$

Suppose that φ_{CN} increases. Clearly, this will affect the right-hand-side of Equation (B.19) only via its effect on the first summand in the squared brackets. Equation (B.16) implies that ${\varphi_{CN}}^{\sigma-1}{w_{CN}}^{1-\sigma}$ increases in response to a raise in φ_{CN} and, hence, that the first summand in the squared brackets in Equation (B.19) decreases. It follows that entering industry i is getting more profitable as φ_{CN} increases.

(ii) If China enters new industries, this will again impact $\frac{\pi_S^{i,v}}{w_S}$ only via its effect on the term in squared brackets in Equation (B.19). This term can be rearranged as follows

$$
\begin{aligned}
& \sum_{\hat{i} \in \mathcal{I}_S \cap \mathcal{I}_{CN}} \frac{\left[\varphi_S \lambda^{\tilde{\mu}_S^{\hat{i}}}\right]^{\sigma-1}}{{\varphi_{CN}}^{\sigma-1}{w_{CN}}^{1-\sigma}\phi + (1-\phi)\left[\varphi_N\right]^{\sigma-1}} + \sum_{\hat{i} \in \mathcal{I}_S \backslash \mathcal{I}_{CN}} \frac{\left[\varphi_S \lambda^{\tilde{\mu}_S^{\hat{i}}}\right]^{\sigma-1}}{(1-\phi)\left[\varphi_N\right]^{\sigma-1}} \\
= & \frac{{\varphi_S}^{\sigma-1}}{(1-\phi){\varphi_N}^{\sigma-1}}\left[\frac{I_{CN}(1-\phi)\left[\varphi_N\right]^{\sigma-1}}{{\varphi_{CN}}^{\sigma-1}{w_{CN}}^{1-\sigma}\phi + (1-\phi)\left[\varphi_N\right]^{\sigma-1}} + (I_{S,-1} - I_{CN}) + \sum_{\hat{i} \in \mathcal{I}_S \backslash \mathcal{I}_{S,-1}} \lambda^{\tilde{\mu}_S^{\hat{i}}(\sigma-1)}\right] ,
\end{aligned} \tag{B.20}
$$

where I_{CN} denotes the number of elements in $\mathcal{I}_{CN}$. Now, Equation (B.16) implies that

$$I_{CN}\frac{{\varphi_{CN}}^{\sigma-1}{w_{CN}}^{1-\sigma}\phi}{{\varphi_{CN}}^{\sigma-1}{w_{CN}}^{1-\sigma}\phi+(1-\phi)\left[\varphi_N\right]^{\sigma-1}}=I_{CN}\left[1-\frac{(1-\phi)\left[\varphi_N\right]^{\sigma-1}}{{\varphi_{CN}}^{\sigma-1}{w_{CN}}^{1-\sigma}\phi+(1-\phi)\left[\varphi_N\right]^{\sigma-1}}\right]$$

increases as I_{CN} goes up. It follows that Expression (B.20) decreases as China enters additional industries which, by (B.19), implies that $\frac{\pi_S^{i,v}}{w_S}$ increases.

(iii) and (iv) Combining Equations (B.17) and (B.18) along with the fact that $\mathcal{I}_S \subseteq \mathcal{I}_{CN}$, we get

$$\frac{\pi_S^{i,v}}{w_S}=\frac{L-\sum_{\hat{i}\in\mathcal{I}_S\setminus\mathcal{I}_{S,-1}}f_S^{\hat{i}}}{\sigma-1}\frac{\left[\lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1}}{\sum_{\hat{i}\in\mathcal{I}_S}\left[\lambda^{\tilde{\mu}_S^{\hat{i}}}\right]^{\sigma-1}}\ . \tag{B.21}$$

Equation (B.21) does not depend on φ_{CN} or the set $\mathcal{I}_{CN}$, which proves that a further growth of China will not impact the growth prospects in the South.

Finally, to see that the growth prospects in the South discretely drop when China leapfrogs, note that the right-hand-side of Equation (B.19) is strictly larger than the right-hand-side of Equation (B.21)

$$\begin{aligned}&\frac{L-\sum_{\hat{i}\in\mathcal{I}_S\setminus\mathcal{I}_{S,-1}}f_S^{\hat{i}}}{\sigma-1}\left[\varphi_S\lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1}\left[(1-\phi)\left[\varphi_N\right]^{\sigma-1}\right]^{-1}\\&\cdot\left[\sum_{\hat{i}\in\mathcal{I}_S\cap\mathcal{I}_{CN}}\frac{\left[\varphi_S\lambda^{\tilde{\mu}_S^{\hat{i}}}\right]^{\sigma-1}}{{\varphi_{CN}}^{\sigma-1}{w_{CN}}^{1-\sigma}\phi+(1-\phi)\left[\varphi_N\right]^{\sigma-1}}+\sum_{\hat{i}\in\mathcal{I}_S\setminus\mathcal{I}_{CN}}\frac{\left[\varphi_S\lambda^{\tilde{\mu}_S^{\hat{i}}}\right]^{\sigma-1}}{(1-\phi)\left[\varphi_N\right]^{\sigma-1}}\right]^{-1}\\>&\frac{L-\sum_{\hat{i}\in\mathcal{I}_S\setminus\mathcal{I}_{S,-1}}f_S^{\hat{i}}}{\sigma-1}\frac{\left[\lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1}}{\sum_{\hat{i}\in\mathcal{I}_S}\left[\lambda^{\tilde{\mu}_S^{\hat{i}}}\right]^{\sigma-1}}\ ,\end{aligned}$$

where the inequality follows from

$${\varphi_{CN}}^{\sigma-1}{w_{CN}}^{1-\sigma}\phi+(1-\phi)\left[\varphi_N\right]^{\sigma-1}>(1-\phi)\left[\varphi_N\right]^{\sigma-1}\ .$$

This completes the proof.

□

C Descriptive Statistics

Table C.1 reports descriptive statistics of our data. In total, we have N=88×140×5=61600 country-industry-period observations. Out of these, about 78% do not meet our presence threshold, leaving us with in 38586 observations for the appearance indicator which, recall, conditions on absence in the base period.[42] In about 5% of these observations, we record the appearance of a new country-industry combination. For the survival indicator, we need to condition on entry, which further allows to use three 5-year windows as base periods only. We, therefore, end up with 1547 observations. On average, 70% of entrants are still present in the subsequent period.

Regarding distance, we report in Table C.1 summarizing statistics for the same subsample without significant exports in the previous period—for if not, the distance is equal to zero by definition. For this reason, a minimum of zero indicates a case where a country has all the occupations required for an industry, but still has no exports. This is consistent with our theory, where a firm in the South may not find it profitable to enter an industry at zero distance if fixed costs of entry are sufficiently large. The distribution of our distance variable is somewhat similar for our two specifications. The average distance $\tilde{\mu}$ indicates that about 8%-9% of qualified employment is missing. There is, however, quite some variation, with a maximum distance of over 70% for both indicators.

Table C.1: Descriptive statistics

Variable	N	mean	sd	min	max
Positive Exports ($x^i_{c,t}$)	61600	0.219	0.414	0.000	1.000
Appearance ($a^i_{c,t}$)	38586	0.050	0.217	0.000	1.000
Survival after entry ($y^i_{c,t}$)	1547	0.696	0.460	0.000	1.000
$\tilde{\mu}^i_{c,t}$ (Cobb-Douglas)	38586	0.079	0.114	0.000	0.718
$\tilde{\mu}^i_{c,t}$ (Leontief)	38586	0.092	0.134	0.000	0.772

Notes: The dataset is obtained converting COMTRADE data to 88 NAICS industries. The benchmark version contains 140 countries and 5 periods (N=88×140×5=61600). $x^i_{c,t}$ is the indicator for presence as described in the main text. Note that for the entry analysis we uses only observations where $x^i_{c,t-1} = 0$ (as we are interested in appearances). This restricts the data to 38586 observations.

Tables C.2 and C.3 show the top-10 and bottom-10 4-digit industries by number of occupations. These lists confirm the large differences across industries in terms of their occupational inputs. The least diversified industry (Leather and Hide Tanning and Finishing) employs only 22 different occupations, while the most diversified industry (Navigational, Measuring, Electromedical, and Control Instruments) uses 207 different occupations—still far less than half of all occupations. One can also observe from the two lists that a ranking of industries by their number of occupations is in line with a qualitative idea of industry 'complexity', with most items in the top-10 (bottom-10) list being relatively high- (low-)

[42]After applying our threshold of $RCA > 1$, and omitting the final 5-year window as a base period because we cannot observe entry in the following period.

tech industries. In the light of these patterns, we can conceive of development as the process of gradually building up the ability to perform additional occupation (occupations), which in turn allows an economy to enter additional, more complex industries.[43]

Table C.2: Top-10 NAICS industries by number of occupations

N of occ.	Naics code	Naics name
207	3345	Navigational, Measuring, Electromedical, and Control Instruments
205	3391	Medical Equipment and Supplies
202	3364	Aerospace Product and Parts
195	3329	Other Fabricated Metal Product
187	3261	Plastics Product
186	3363	Motor Vehicle Parts
185	3254	Pharmaceutical and Medicine
184	3323	Architectural and Structural Metals
182	3339	Other General Purpose Machinery
174	2111	Oil and Gas Extraction

Table C.3: Bottom-10 NAICS industries by number of occupations

N of occ.	Naics code	Naics name
22	3161	Leather and Hide Tanning and Finishing
36	3151	Apparel Knitting Mills
40	3169	Other Leather and Allied Product
41	3159	Apparel Accessories and Other Apparel
44	3162	Footwear
46	3122	Tobacco
47	3131	Fiber, Yarn, and Thread Mills
48	3274	Lime and Gypsum Product
68	3133	Textile and Fabric Finishing and Fabric Coating Mills
69	3117	Seafood Product Preparation and Packaging

Notes: Source: Occupational Employment Statistics 2016 (Bureau of Labor Statistics).

[43] Schetter (2019) shows that along the convergence path, countries develop an export basket that is increasingly similar to the frontier. They do this by increasing their exports in complex products (Hidalgo and Hausmann, 2009; Schetter, 2019).

References

Acemoglu, D. (2003). Patterns of skill premia. *Review of Economic Studies*, 70(2):199–230.

Acemoglu, D., Aghion, P., and Zilibotti, F. (2006). Distance to frontier, selection, and economic growth. *Journal of the European Economic Association*, 4(1):37–74.

Acemoglu, D. and Guerrieri, V. (2008). Capital deepening and nonbalanced economic growth. *Journal of Political Economy*, 116(3):467–498.

Acemoglu, D. and Restrepo, P. (2018). The race between man and machine: Implications of technology for growth, factor shares, and employment. *American Economic Review*, 108(6):1488–1542.

Aghion, P. and Howitt, P. (1992). A model of growth through creative destruction. *Econometrica*, 60(2):323–351.

Aghion, P., Howitt, P., and Mayer-Foulkes, D. (2005). The effect of financial development on convergence: Theory and evidence. *The Quarterly Journal of Economics*, 120(1):173–222.

Alder, S., Boppart, T., and Müller, A. (2022). A theory of structural change that can fit the data. *American Economic Journal: Macroeconomics*, 14(2):160–206.

Allen, T. and Donaldson, D. (2020). Persistence and path dependence in the spatial economy. Working Paper 28059, NBER.

Anderson, J. E. (1979). A theoretical foundation for the gravity equation. *The American Economic Review*, 69(1):106–116.

Arkolakis, C., Ramondo, N., Rodríguez-Clare, A., and Yeaple, S. (2018). Innovation and production in the global economy. *American Economic Review*, 108(8):2128–2173.

Armington, P. S. (1969). A theory of demand for products distinguished by place of production. *Staff Papers - International Monetary Fund*, 16(1):159.

Atkin, D., Costinot, A., and Fukui, M. (2021). Globalization and the ladder of development: Pushed to the top or held at the bottom? Working Paper 29500, NBER. Series: Working Paper Series.

Autor, D., Dorn, D., Hanson, G. H., Pisano, G., and Shu, P. (2020). Foreign competition and domestic innovation: Evidence from US patents. *American Economic Review: Insights*, 2(3):357–74.

Bahar, D., Hauptmann, A., Özgüzel, C., and Rapoport, H. (2022). Migration and knowledge diffusion: The effect of returning refugees on export performance in the former Yugoslavia. *The Review of Economics and Statistics*. https://doi.org/10.1162/rest_a_01165.

Bahar, D. and Rapoport, H. (2018). Migration, knowledge diffusion and the comparative advantage of nations. *The Economic Journal*, 128(612):F273–F305.

Balassa, B. (1965). Trade liberalisation and 'revealed' comparative advantage. *The Manchester School*, 33(2):99–123.

Balland, P.-A., Broekel, T., Diodato, D., Giuliani, E., Hausmann, R., O'Clery, N., and Rigby, D. (2022). The new paradigm of economic complexity. *Research Policy*, 51(3):104450.

Barro, R. J. and Sala-I-Martin, X. (1992). Convergence. *Journal of Political Economy*, 100(2):223–251.

Barro, R. J., Sala-I-Martin, X., Blanchard, O. J., and Hall, R. E. (1991). Convergence across states and regions. *Brookings Papers on Economic Activity*, 1991(1):107–182.

Bartelme, D., Costinot, A., Donaldson, D., and Rodriguez-Clare, A. (2019). The textbook case for industrial policy: Theory meets data. Working Paper 26193, NBER.

Baumol, W. J. (1967). Macroeconomics of unbalanced growth: The anatomy of urban crisis. *The American Economic Review*, 57(3):415–426.

Blonigen, B. A. (2016). Industrial policy and downstream export performance. *The Economic Journal*, 126(595):1635–1659.

Bloom, N., Draca, M., and Van Reenen, J. (2015). Trade induced technical change? The impact of Chinese imports on innovation, IT and productivity. *The Review of Economic Studies*, 83(1):87–117.

Bloom, N., Romer, P., Terry, S. J., and Van Reenen, J. (2020). Trapped factors and China's impact on global growth. *The Economic Journal*, forthcoming.

Boehm, J., Dhingra, S., and Morrow, J. (2022). The comparative advantage of firms. *Journal of Political Economy*. https://doi.org/10.1086/720630.

Boppart, T. (2014). Structural change and the Kaldor facts in a growth model with relative price effects and non-Gorman preferences. *Econometrica*, 82(6):2167–2196.

Brummitt, C. D., Gómez-Liévano, A., Hausmann, R., and Bonds, M. H. (2020). Machine-learned patterns suggest that diversification drives economic development. *Journal of*

The Royal Society Interface, 17(162):20190283.

Buera, F. J., Hopenhayn, H., Shin, Y., and Trachter, N. (2021). Big push in distorted economies. Working Paper 28561, NBER.

Buera, F. J. and Kaboski, J. P. (2012). The rise of the service economy. *American Economic Review*, 102(6):2540–2569.

Buera, F. J., Kaboski, J. P., Rogerson, R., and Vizcaino, J. I. (2022). Skill-biased structural change. *Review of Economic Studies*, 89(2):592–625.

Buera, F. J. and Oberfield, E. (2020). The global diffusion of ideas. *Econometrica*, 88(1):83–114.

Bustos, S., Gomez, C., Hausmann, R., and Hidalgo, C. A. (2012). The dynamics of nestedness predicts the evolution of industrial ecosystems. *PLOS ONE*, 7(11):e49393.

Cadot, O., Carrère, C., and Strauss-Kahn, V. (2010). Export diversification: What's behind the hump? *Review of Economics and Statistics*, 93(2):590–605.

Cadot, O., Carrere, C., and Strauss-Kahn, V. (2013). Trade diversification, income, and growth: What do we know? *Journal of Economic Surveys*, 27(4):790–812.

Caselli, F. and Coleman II, W. J. (2001). The U.S. structural transformation and regional convergence: A reinterpretation. *Journal of Political Economy*, 109(3):584–616.

Chenery, H. B. (1960). Patterns of industrial growth. *American Economic Review*, 50(4):624–654.

Choi, J. and Levchenko, A. A. (2021). The long-term effects of industrial policy. Working Paper 29263, NBER.

Choi, J. and Shim, Y. (2022). Technology adoption and late industrialization. Mimeo.

Chu, A. C., Cozzi, G., Lai, C.-C., and Liao, C.-H. (2015). Inflation, R&D and growth in an open economy. *Journal of International Economics*, 96(2):360–374.

Ciccone, A. (2002). Input chains and industrialization. *Review of Economic Studies*, 69(3):565–587.

Ciccone, A. and Papaioannou, E. (2009). Human capital, the structure of production, and growth. *Review of Economics and Statistics*, 91(1):66–82.

De Benedictis, L., Gallegati, M., and Tamberi, M. (2009). Overall trade specialization and economic development: Countries diversify. *Review of World Economics*, 145(1):37–55.

Desmet, K. and Rossi-Hansberg, E. (2014). Spatial development. *American Economic*

Review, 104(4):1211–1243.

Diodato, D., Hausmann, R., and Neffke, F. (2020). The impact of return migration from the U.S. on employment and wages in Mexican cities. Papers in Evolutionary Economic Geography 20.12, Utrecht University.

Duarte, M. and Restuccia, D. (2020). Relative prices and sectoral productivity. *Journal of the European Economic Association*, 18(3):1400–1443.

Duernecker, G., Herrendorf, B., and Valentinyi, Á. (2021). Structural change within the services sector and the future of cost disease. Mimeo.

Eaton, J. and Kortum, S. (2002). Technology, geography, and trade. *Econometrica*, 70(5):1741–1779.

Eckert, F. and Peters, M. (2018). Spatial structural change. Mimeo.

Ellison, G., Glaeser, E., and Kerr, W. (2010). What causes industry agglomeration? Evidence from coagglomeration patterns. *American Economic Review*, 100(3):1195–1213.

Epifani, P. and Gancia, G. (2011). Trade, markup heterogeneity and misallocations. *Journal of International Economics*, 83(1):1–13.

Foellmi, R., Hanslin Grossmann, S., and Kohler, A. (2018). A dynamic North-South model of demand-induced product cycles. *Journal of International Economics*, 110:63–86.

Foellmi, R. and Zweimüller, J. (2008). Structural change, Engel's consumption cycles and Kaldor's facts of economic growth. *Journal of Monetary Economics*, 55(7):1317–1328.

Galor, O. and Mountford, A. (2008). Trading population for productivity: Theory and evidence. *Review of Economic Studies*, 75(4):1143–1179.

Gersbach, H., Schetter, U., and Schmassmann, S. (2019). From local to global: A unified theory of public basic research. Discussion Paper 13833, CEPR.

Gersbach, H., Schneider, M. T., and Schneller, O. (2013). Basic research, openness, and convergence. *Journal of Economic Growth*, 18(1):33–68.

Gollin, D., Parente, S., and Rogerson, R. (2002). The role of agriculture in development. *American Economic Review*, 92(2):160–164.

Grossman, G. M. and Helpman, E. (1991a). Endogenous product cycles. *The Economic Journal*, 101(408):1214–1229.

Grossman, G. M. and Helpman, E. (1991b). *Innovation and Growth in the Global Econ-*

omy. MIT Press, Cambridge, MA.

Grossman, G. M. and Helpman, E. (1991c). Quality ladders in the theory of growth. *The Review of Economic Studies*, 58(1):43–61.

Grossman, G. M. and Helpman, E. (1995). Chapter 25–technology and trade. In Gene M. Grossman and Kenneth Rogoff, editor, *Handbook of International Economics*, volume Volume 3, pages 1279–1337. Elsevier.

Harrison, A. and Rodríguez-Clare, A. (2010). Chapter 63 - Trade, foreign investment, and industrial policy for developing countries. In Rodrik, D. and Rosenzweig, M., editors, *Handbook of Development Economics*, volume 5, pages 4039–4214. Elsevier.

Hausmann, R. and Hidalgo, C. A. (2011). The network structure of economic output. *Journal of Economic Growth*, 16(4):309–342.

Hausmann, R., Hidalgo, C. A., Bustos, S., Coscia, M., Chung, S., Jimenez, J., Simoes, A., and Yildirim, M. A. (2011). *The Atlas of Economic Complexity: Mapping Paths to Prosperity*. https://atlas.media.mit.edu/atlas/.

Hausmann, R. and Klinger, B. (2006). Structural transformation and patterns of comparative advantage in the product space. Working Paper 128, CID at Harvard University.

Hausmann, R. and Klinger, B. (2007). The structure of the product space and the evolution of comparative advantage. Working Paper 146, CID at Harvard University.

Hausmann, R. and Neffke, F. M. H. (2019). The workforce of pioneer plants: The role of worker mobility in the diffusion of industries. *Research Policy*, 48(3):628–648.

Hausmann, R. and Rodrik, D. (2003). Economic development as self-discovery. *Journal of Development Economics*, 72(2):603–633.

Hausmann, R. and Rodrik, D. (2006). Doomed to choose: Industrial policy as predicament. Mimeo, Harvard Kennedy School.

Hicks, J. R. (1953). An inaugural lecture. *Oxford Economic Papers*, 5(2):117–135.

Hidalgo, C. A., Balland, P.-A., Boschma, R., Delgado, M., Feldman, M., Frenken, K., Glaeser, E., He, C., Kogler, D. F., Morrison, A., Neffke, F., Rigby, D., Stern, S., Zheng, S., and Zhu, S. (2018). The principle of relatedness. In Morales, A. J., Gershenson, C., Braha, D., Minai, A. A., and Bar-Yam, Y., editors, *Unifying Themes in Complex Systems IX*, Springer Proceedings in Complexity, pages 451–457, Cham. Springer International Publishing.

Hidalgo, C. A. and Hausmann, R. (2009). The building blocks of economic complexity.

Proceedings of the National Academy of Sciences, 106(26):10570–10575.

Hidalgo, C. A., Klinger, B., Barabási, A.-L., and Hausmann, R. (2007). The product space conditions the development of nations. *Science*, 317(5837):482–487.

Howitt, P. (2000). Endogenous growth and cross-country income differences. *American Economic Review*, 90(4):829–846.

Hsieh, C.-T. (2002). What explains the industrial revolution in east Asia? Evidence from the factor markets. *American Economic Review*, 92(3):502–526.

Imbs, J. and Wacziarg, R. (2003). Stages of diversification. *American Economic Review*, 93(1):63–86.

Johnson, W. (2020). Economic growth and the evolution of comparative advantage in a network of industries. Mimeo, Dartmouth College.

Jones, C. I. (2011). Intermediate goods and weak links in the theory of economic development. *American Economic Journal: Macroeconomics*, 3(2):1–28.

Jäggi, A., Schetter, U., and Schneider, M. T. (2021). Inequality, openness, and growth through creative destruction. Research Fellow and Graduate Student Working 130, CID at Harvard University. https://dx.doi.org/10.2139/ssrn.3795632.

Kehoe, T. J. and Ruhl, K. J. (2013). How important is the new goods margin in international trade? *Journal of Political Economy*, 121(2):358–392.

Kongsamut, P., Rebelo, S., and Xie, D. (2001). Beyond balanced growth. *Review of Economic Studies*, 68(4):869–882.

Kremer, M. (1993). The O-ring theory of economic development. *Quarterly Journal of Economics*, 108(3):551–575.

Krugman, P. (1979). A model of innovation, technology transfer, and the world distribution of income. *Journal of Political Economy*, 87(2):253–266.

Kuznets, S. (1957). Quantitative aspects of the economic growth of nations: II. Industrial distribution of national product and labor force. *Economic Development and Cultural Change*, 5(4):1–111.

Lane, N. (2020). The new empirics of industrial policy. *Journal of Industry, Competition and Trade*, 1(2):1–26.

Liu, E. (2019). Industrial policies in production networks. *Quarterly Journal of Economics*, 134(4):1883–1948.

Lucas, R. E. (1993). Making a miracle. *Econometrica*, 61(2):251–272.

Matsuyama, K. (1992). Agricultural productivity, comparative advantage, and economic growth. *Journal of Economic Theory*, 58(2):317–334.

Matsuyama, K. (1995). Complementarities and cumulative processes in models of monopolistic competition. *Journal of Economic Literature*, 33(2):701–729.

Matsuyama, K. (2000). A Ricardian model with a continuum of goods under nonhomothetic preferences: Demand complementarities, income distribution, and North-South trade. *Journal of Political Economy*, 108(6):1093–1120.

Matsuyama, K. (2019). Engel's law in the global economy: Demand-induced patterns of structural change, innovation, and trade. *Econometrica*, 87(2):497–528.

McNerney, J., Savoie, C., Caravelli, F., Carvalho, V. M., and Farmer, J. D. (2022). How production networks amplify economic growth. *Proceedings of the National Academy of Sciences*, 119(1):e2106031118.

Melitz, M. J. (2003). The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica*, 71(6):1695–1725.

Neffke, F. and Henning, M. (2013). Skill relatedness and firm diversification. *Strategic Management Journal*, 34(3):297–316.

Ngai, L. R. and Pissarides, C. A. (2007). Structural change in a multisector model of growth. *American Economic Review*, 97(1):429–443.

Nunn, N. and Trefler, D. (2010). The structure of tariffs and long-term growth. *American Economic Journal: Macroeconomics*, 2(4):158–194.

Ottinger, S. (2020). Immigrants, Industries, and Path Dependence. Mimeo.

Parteka, A. and Tamberi, M. (2013). What determines export diversification in the development process? Empirical assessment. *The World Economy*, 36(6):807–826.

Peters, M. and Zilibotti, F. (2021). Creative destruction, distance to frontier, and economic development. Working Paper 29333, NBER.

Rodrik, D. (2012). Unconditional convergence in manufacturing. *Quarterly Journal of Economics*, 128(1):165–204.

Romer, P. M. (1987). Growth based on increasing returns due to specialization. *American Economic Review*, 77(2):56–62.

Romer, P. M. (1990). Endogenous technological change. *Journal of Political Economy*, 98(5):S71–S102.

Sampson, T. (2016). Dynamic selection: An idea flows theory of entry, trade, and growth. *The Quarterly Journal of Economics*, 131(1):315–380.

Schetter, U. (2019). A structural ranking of economic complexity. Research Fellow and Graduate Student Working Paper 119, CID at Harvard University. https://dx.doi.org/10.2139/ssrn.3485842.

Schetter, U. (2020). Quality differentiation, comparative advantage, and international specialization across products. Research Fellow and Graduate Student Working Paper 126, CID at Harvard University. http://dx.doi.org/10.2139/ssrn.3091581.

Stokey, N. L. (1991). The volume and composition of trade between rich and poor countries. *Review of Economic Studies*, 58(1):63–80.

Sutton, J. and Trefler, D. (2016). Capabilities, wealth, and trade. *Journal of Political Economy*, 124(3):826–878.

van Dam, A. and Frenken, K. (2020). Variety, complexity and economic development. *Research Policy*, page 103949.

Vernon, R. (1966). International investment and international trade in the product cycle. *The Quarterly Journal of Economics*, 80(2):190–207.

Supplementary Material

S1 Mathematical Appendix

In this appendix, we provide further details on some of the discussions in the main text.

S1.1 Further Details on Entry with a Continuum of Firms

In this appendix we consider the case of a continuum of measure $\frac{1}{1+n}$ of firms by country and industry, each equipped with a distinct variety. As before, firms can freely enter and entry involves a fixed cost. In turn, this implies that our analysis of the main text directly applies with the only minor caveat that a subset of firms only may find it profitable to enter. This would complicate the exposition without adding anything of substance to our main arguments, and we therefore consider the case of one firm by country and industry instead. In fact if fixed cost have an aggregate component that is sufficiently important and / or the impact of entry on wages in the South is not too big, it is profitable for either all or no firm to enter an industry in the South, and our analysis from the main text directly applies. We will get back to this point at the end of this section.

Things are different, however, if we allow the fixed cost to be partly aggregate, i.e., if we allow the per-firm fixed cost to decline with the mass of entrants. In such case, a strictly positive mass of entrants is required for entry to be profitable, giving rise to coordination problems as discussed in Section 5.2.

In particular, let us assume that the fixed cost borne by an individual entrant in the South and industry i are

$$\tilde{f}_S^i(\gamma_S^i) := \frac{f_S^i}{(\gamma_S^i)^\delta} \ , \quad \delta \in [0,1] \tag{S1.1}$$

units of labor, where δ is a parameter capturing the public nature of fixed cost and $0 \leq \gamma_S^i$ is the share of all firms that enter. $\delta = 0$ and hence $\tilde{f}_S^i(\gamma_S^i) = f_S^i$ corresponds to the case of pure private fixed cost, while $\delta = 1$ and hence $\tilde{f}_S^i(\gamma_S^i) = \frac{f_S^i}{\gamma_S^i}$ corresponds to the case of pure 'public' fixed cost of entry. Suppose, for simplicity, that in the current period there is entry in one industry only, namely industry $\hat{\imath}$. Derivations analogous to those for Equation (B.2) then imply that entry is profitable if

$$\frac{{\gamma_S^{\hat{\imath}}}^{\delta} L - f_S^{\hat{\imath}} \gamma_S^{\hat{\imath}}}{(\sigma-1)\left[I_{S,-1}/\lambda^{\tilde{\mu}_S^{\hat{\imath}}(\sigma-1)} + \gamma_S^{\hat{\imath}}\right]} \geq f_S^{\hat{\imath}} \ ,$$

which we can rearrange to

$$\gamma_S^{\hat{i}\,\delta} L - f_S^{\hat{i}} \gamma_S^{\hat{i}} \sigma \geq f_S^{\hat{i}} (\sigma - 1) \frac{I_{S,-1}}{\lambda^{\tilde{\mu}_S^{\hat{i}}(\sigma-1)}} \ . \tag{S1.2}$$

Clearly, for any $\delta > 0$ and as long as $f_S^{\hat{i}} > 0$ there is a positive mass of firms needed in order for entry to be profitable, potentially giving rise to coordination failures as discussed in Section 5.2.

In the main text, we consider the case of one firm by country and industry. This is analogous to a setup with a continuum of firms if in the latter case either all or no firm enters. A sufficient condition for this is if the left-hand-side of Condition (S1.2) is increasing in γ_S^i. Differentiating the left-hand-side with respect to $\gamma_S^{\hat{i}}$ and rearranging terms reveals that this is the case whenever

$$\gamma_S^{\hat{i}\,\delta-1} \geq \frac{f_S^{\hat{i}} \sigma}{\delta L} \ . \tag{S1.3}$$

Hence, if the right-hand-side of Condition (S1.3) is smaller than 1 indeed either all or no firm enter a new industry in the South.

S1.2 Robustness of Proposition 2

In this appendix, we provide further details on the robustness of Proposition 2 with regards to our simplifying assumptions.

S1.2.1 Overview

To simplify the exposition, we assume that agents are myopic in the baseline version of our theoretical model. The main implication of this assumption is that firms compare current profits with the fixed cost of entry when deciding on whether or not to enter. This is obviously a simplification and we may expect firms to be forward looking in their entry decision. Note, however, that industries are perfectly symmetric after the entry period and, in particular, the world trade share—and therefore total exports of the South—is the same for all industries $i \in \mathcal{I}_{S,-1}$ (see Equation (14)). Hence, as long as entering firms are not anticipating the effect of their entry on future entry in other industries, forward looking behavior on behalf of the firms does not change the qualitative statements in Proposition 2.[44]

[44] As we discuss in Section 5.1, industries may differ in their potential to facilitate entry in other industries in future. Future entry in new industries increases w_S and, hence, decreases profits (see Equation (15)). It would therefore impact today's entry decision, if anticipated. It is arguably unrealistic

To simplify the discussions, we have further assumed that the South is small and that it is embedded in a world with many perfectly symmetric countries, all at the frontier. In Section 5.3, we show that the South is also more likely to enter nearby industries if parts of the rest-of-the-world are not at the frontier.[45] In Section S1.2.2 below, we show that this is also the case when the South is no longer small vis-à-vis the rest-of-the-world. The key difference is that in such case entry in the South impacts the aggregate price level in an industry as well as aggregate demand. Yet, the South is still more likely to enter nearby industries.

We further assumed that fixed cost are in terms of labor and occur in the entry period only. We assumed zero fixed cost in future periods to avoid the need of keeping track of potential exit. This simplifies the exposition but is not essential for our main insights.[46] The assumption of fixed cost in terms of labor implies that entry gets increasingly costly as the South develops. As an alternative, we could assume that fixed cost are in terms of the final consumption aggregator and, still, the South is more likely to enter nearby industries.[47]

Entry into a new industry may involve uncertainty, which is absent from the baseline version of our model. It is therefore interesting to note that Proposition 2 is also robust to introducing uncertainty, as shown in Section 4 and Appendix A.

Finally, we assume that there are no comparative advantages beyond those at the extensive margin if the South is not active in an industry. Again, this is for simplicity only and,

that firms anticipate such general equilibrium effects on otherwise unrelated industries, and we therefore do not believe them to be a major caveat in practice. We note that if firms anticipated such effects it would make income traps as discussed in Section 5.1 even more likely, as a firm's future profits are ceteris paribus highest without future growth in the South.

[45] Of course, industries differ in their competitiveness if not all countries in the rest-of-the world are at the frontier and fully diversified, and such differences feed back into the profit potential for entrants in the South. Yet, while the competitive environment obviously matters for entry in the South—as does the industry size α, for example—regarding our above insights it merely implies that we need to condition our statements on this competitive environment.

[46] With myopic entry, future exit will trivially not affect the order of industries in terms of their probability of entry in the South. This, however, is also true if firms are concerned with future exit, because—as already noted above—total sales and, hence, variable profits are the same across all industries $i \in \mathcal{I}_{S,-1}$.

[47] In fact, the countervailing effect of a larger wage increase when entering nearby industries is smaller in this case as it does not affect the fixed cost of entry. Note that it is nonetheless the case that entry in different industries are strategic substitutes, i.e., the sort of multiple equilibria that form the basis for Proposition 4 can still arise. To see this, observe from Equation (15) that with fixed cost in terms of the final consumption aggregator, the representative firm in industry i finds it beneficial to enter if

$$\pi_S^{i,v} = \left[\varphi_S \lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1} {w_S}^{1-\sigma} \left[\varphi_N\right]^{1-\sigma} \frac{\alpha L}{\sigma - 1} \geq f_S^i \ ,$$

and the result follows from the fact that w_S increases as the South enters new industries.

in general, such comparative advantages would simply require conditioning our results on them. Nevertheless, it is interesting to note that all our results also apply to a variant of our model with an O-ring production process and an endogenous choice of quality following Kremer (1993); Schetter (2020). The O-ring process provides an intuitive narrative for the network structure of industries, and it provides a micro-foundation for why there are no comparative advantages across industries that differ largely in their complexity—see Fact 1 and Tables C.2 and C.3. Moreover, it admits interpretations as both a Leontief or a Cobb-Douglas production function. And it can rationalize Fact 3 when using log distance (instead of levels) as a control. Our empirical finding is robust to this choice—see Table S1. Further details are available upon request.

S1.2.2 Details on Entry in a Large South

In this appendix, we show that the South is more likely to enter nearby industries also if it is large, i.e., if it is embedded in a world with a finite number n of countries in the North.

With a finite number of countries in the North, the wage in the South satisfies (see Appendix B.1)

$$\left[nL + w_s\left(L - \sum_{i\in\mathcal{I}_S\setminus\mathcal{I}_{S,-1}} f_S^i\right)\right]\sum_{i\in\mathcal{I}_S}\frac{\alpha\left[\varphi_S\lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1}{w_S}^{-\sigma}}{\left[\varphi_S\lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1}{w_S}^{1-\sigma}+n\left[\varphi_N\right]^{\sigma-1}} = L - \sum_{i\in\mathcal{I}_S\setminus\mathcal{I}_{S,-1}} f_S^i\ , \tag{S1.4}$$

and the firm in industry i in the South finds it profitable to enter if[48]

$$\frac{\pi_S^{i,v}}{w_S} = \frac{\left[\varphi_S\lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1}{w_S}^{-\sigma}\alpha\left[nL + w_S\left(L - \sum_{\hat{i}\in\mathcal{I}_S\setminus\mathcal{I}_{S,-1}} f_S^{\hat{i}}\right)\right]}{\left[\left[\varphi_S\lambda^{\tilde{\mu}_S^i}\right]^{\sigma-1}{w_S}^{1-\sigma}+n\left[\varphi_N\right]^{\sigma-1}\right](\sigma-1)} \geq f_S^i\ . \tag{S1.5}$$

We need to show that for any pair of industries, i_1, i_2 such that $\tilde{\mu}_S^{i_1} < \tilde{\mu}_S^{i_2}$, and a given fixed cost $f_S^{i_1} = f_S^{i_2} = f$, $\frac{\pi_S^{i,v}}{w_S}$ is larger when entering industry i_1 than when entering industry i_2. To do so, we proceed in three steps. We first show that $w_S(i_1) > w_S(i_2)$, where—with a slight abuse of notation—we use here and below $w_S(i)$ to denote the wage that would—all else equal—prevail in the South when entering industry i. We then show that $\left[\lambda^{\tilde{\mu}_S^{i_1}}\right]^{\sigma-1} w_S(i_1)^{1-\sigma} > \left[\lambda^{\tilde{\mu}_S^{i_2}}\right]^{\sigma-1} w_S(i_2)^{1-\sigma}$. We finally show the desired result.

[48] The expression for profits of the representative firm in the South in industry i given in Equation (S1.5) assumes that the firm has no effect on aggregate prices in industry i. With a finite number of countries and one firm per country and industry this is no longer the case. Note, however, that we can interpret our analysis as one where there is a continuum of measure $\frac{1}{n+1}$ of firms in each country and industry, in which case an individual firm no longer has an influence on aggregate prices. See Supplementary Material S1.1.

Step 1: $w_S(i_1) > w_S(i_2)$

We proceed by contradiction. Suppose that $w_S(i_1) \leq w_S(i_2)$. Then the LHS of Equation (S1.4) is strictly larger when entering industry i_1 than when entering industry i_2, a contradiction to $w_S(i_1)$ and $w_S(i_2)$ being the equilibrium wages when entering industry i_1 and i_2, respectively.

Step 2: $\left[\lambda^{\tilde{\mu}_S^{i_1}}\right]^{\sigma-1} w_S(i_1)^{1-\sigma} > \left[\lambda^{\tilde{\mu}_S^{i_2}}\right]^{\sigma-1} w_S(i_2)^{1-\sigma}$

We proceed by contradiction. Suppose that

$$\left[\lambda^{\tilde{\mu}_S^{i_1}}\right]^{\sigma-1} w_S(i_1)^{1-\sigma} \leq \left[\lambda^{\tilde{\mu}_S^{i_2}}\right]^{\sigma-1} w_S(i_2)^{1-\sigma} .$$

Then step 1 implies that the LHS of Equation (S1.4) is strictly smaller when entering industry i_1 than when entering industry i_2, a contradiction to $w_S(i_1)$ and $w_S(i_2)$ being the equilibrium wages when entering industry i_1 and i_2, respectively.

Step 3:

Solving Equation (S1.4) for $\left[\frac{nL}{w_S} + \left(L - \sum_{i\in\mathcal{I}_S\setminus\mathcal{I}_{S,-1}} f_S^i\right)\right]$, plugging it into Equation (S1.5), and rearranging terms, we get for $\hat{i} \in \{i_1, i_2\}$

$$\frac{\pi_S^{\hat{i},v}}{w_S(\hat{i})} = \frac{L - \sum_{i\in\mathcal{I}_S\setminus\mathcal{I}_{S,-1}} f_S^i}{\sigma-1} \frac{\left[\varphi_S \lambda^{\tilde{\mu}_S^{\hat{i}}}\right]^{\sigma-1} w_S(\hat{i})^{1-\sigma}}{\left[\varphi_S \lambda^{\tilde{\mu}_S^{\hat{i}}}\right]^{\sigma-1} w_S(\hat{i})^{1-\sigma} + n\left[\varphi_N\right]^{\sigma-1}}$$
$$\left[\sum_{i\in\mathcal{I}_S} \frac{\left[\varphi_S \lambda^{\tilde{\mu}_S^{i}}\right]^{\sigma-1} w_S(\hat{i})^{1-\sigma}}{\left[\varphi_S \lambda^{\tilde{\mu}_S^{i}}\right]^{\sigma-1} w_S(\hat{i})^{1-\sigma} + n\left[\varphi_N\right]^{\sigma-1}}\right]^{-1} .$$

The second fraction is larger for i_1 than for i_2 by step 2. The term inside squared brackets is smaller for i_1 than for i_2 by step 1. It follows that $\frac{\pi_S^{\hat{i},v}}{w_S(\hat{i})}$ is larger when entering industry i_1 than when entering industry i_2, which shows the desired result.

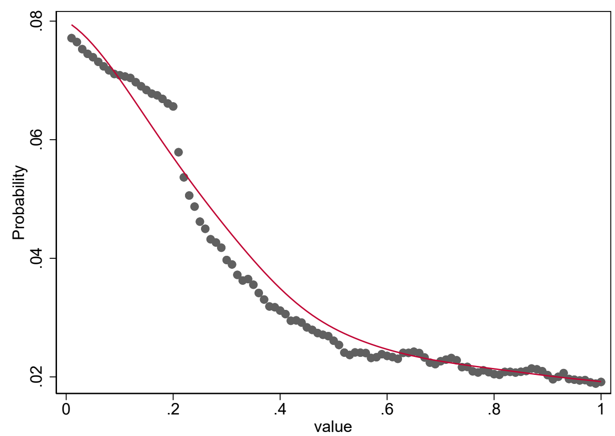
S2 Empirical Appendix

This appendix presents robustness checks for our main regressions and further details on our data.

S2.1 Robustness of Fact 3

This appendix presents robustness checks for our main motivating fact, Fact 3.

Figure S1: Probability of industry appearance and distance (measured with employment shares)



Notes: The horizontal axis represents percentiles of $\tilde{\mu}^i_{c,t}$ (p/100). The vertical axis is a moving average of appearance ($a^i_{c,t}$, as in Equation (2)) for an interval ($\pm.2$) around the corresponding x-axis value. The trend line depicts a LOWESS smooth.

Figure S1 replicates Figure 3, measuring distance based on employment instead of wage-bill shares. With this alternative choice, we can map our empirical framework to a Leontief production function instead of a Cobb-Douglas.

Table S1 reports robustness checks for the main results in Table 1. In our benchmark analysis, we consider a country-industry combination as present if it has an RCA>1 for at least three out of the five years in a period. The second and third row of Table S1 show that our results are robust to the use of a lower relative threshold (RCA>0.5) as well as to the use of a fixed absolute threshold (10M USD of exports).

Next, in the benchmark we measure distance ($\tilde{\mu}^i_{c,t}$) in levels, which is the functional form suggested by our theory—see Section 3.3. The fourth row (‘$distance^i_{c,t} = \ln(\tilde{\mu}^i_{c,t})$’) shows that Fact 3 is robust to using log-distance instead.

In the baseline specification, we use data on occupational inputs by industry from the US Occupational Employment Statistics for 2002. This is a mid-point in our sample period. Rows 5 and 6 of Table S1 show that our results are robust to using US Occupational Employment Statistics from 2016 (OES2016) or industry-occupation data from Mexico instead (obtained aggregating ENOE survey data for the year 2005).

Rows 7 and 8 test for alternative country (‘*Aggregation of countries*’) and industry aggregations (‘*Alternative concordance*’). See the discussion in Supplementary Material S2.5 for further details.

Table S1: Robustness: Probability of industry appearance and distance

	(1)	(2)	(3)	(4)	(5)
Benchmark	-0.099***	-0.202***	-0.052**	-0.062***	-0.092***
	(0.015)	(0.018)	(0.020)	(0.018)	(0.020)
Alternative RCA threshold (0.5)	-0.138***	-0.350***	-0.128***	-0.067***	-0.093***
	(0.021)	(0.029)	(0.041)	(0.025)	(0.024)
Fixed threshold (10M USD)	-0.041***	-0.178***	-0.094***	-0.048***	-0.101***
	(0.016)	(0.021)	(0.026)	(0.014)	(0.024)
$distance^i_{c,t} = \ln(\tilde{\mu}^i_{c,t})$	-0.009***	-0.013***	-0.005***	-0.008***	-0.011***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)
Tech. from OES2016	-0.101***	-0.209***	-0.071***	-0.089***	-0.121***
	(0.014)	(0.016)	(0.020)	(0.016)	(0.016)
Tech. from Mexico (ENOE)	-0.004***	-0.010***	-0.005	-0.003	-0.006***
	(0.001)	(0.002)	(0.003)	(0.002)	(0.002)
Aggregation of countries	-0.096***	-0.196***	-0.054***	-0.060***	-0.089***
	(0.015)	(0.017)	(0.021)	(0.018)	(0.020)
Alternative concordance	-0.108***	-0.213***	-0.084***	-0.076***	-0.114***
	(0.014)	(0.018)	(0.017)	(0.017)	(0.019)
Probit	-1.562***	-3.154***		-1.095***	-1.650***
	(0.256)	(0.331)		(0.299)	(0.323)
Dummies	ct	it	ci	c,i	ct,it

Notes: The dependent variable is appearance of an industry in t ($a^i_{c,t}$), as defined in Equation (2). The independent variable is $\tilde{\mu}^i_{c,t}$ (Cobb-Douglas), that is, using the wage-bill share for μ^i_τ. Rows indicate the following robustness checks: *'Benchmark'* is the reference, as in the main text; *'Alternative RCA'* uses 0.5 instead of 1 as the threshold; *'Fixed threshold'* requires a minimum of 10M USD to consider a country an exporter of a commodity; *'$distance^i_{c,t} = \ln(\tilde{\mu}^i_{c,t})$'* uses the logarithm of our main control; *'Tech from OES2016'* uses BLS's occupational employment statistics from 2016 to compute μ^i_τ; *'Tech from Mexico (ENOE)'* uses a dataset from Mexico (Encuesta Nacional de Ocupación y Empleo) to compute μ^i_τ; *'Aggregation of countries'* drops the small countries in the first column of Table S3; *'Alternative concordance'* uses an alternative method to convert HS to NAICS, as described in Supplementary Material S2.5.1; *'Probit'* uses a probit estimator. Note that there are too many dummies in model (3) for this estimator. Country-level cluster robust standard errors in parentheses. Significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Lastly, we estimate the coefficient using a probit estimator. As we discuss in Section 3.3, the linear probability model can be justified by assuming that $\ln(f^i_S)$ is uniformly distributed. An alternative assumption of normal distribution implies, instead, the probit model.

All in all, the robustness analysis confirms the negative and significant relationship between distance and appearance. Notably, this is always the case in column (5), where we include the dummies suggested by our theory.

S2.2 Robustness of Table 2

Table S2 replicates the robustness checks of Appendix S2.1 for Table 2. Not surprisingly, considering the much smaller sample size (the sample is conditioned on entry), this result is somewhat less robust. Nevertheless, Table S2 broadly confirms the findings from Table 2.

Table S2: Robustness: Probability of survival and distance

	(1)	(2)	(3)	(4)	(5)
Benchmark	-0.092	-0.957***	-3.143***	-0.448***	-0.479**
	(0.206)	(0.210)	(0.970)	(0.171)	(0.235)
Alternative RCA threshold (0.5)	0.083	-0.994***	-1.356	0.159	0.019
	(0.205)	(0.253)	(0.904)	(0.254)	(0.265)
Fixed threshold (10M USD)	-0.327***	-0.164**	-1.348***	0.003	-0.131
	(0.125)	(0.071)	(0.359)	(0.062)	(0.167)
$distance^i_{c,t} = \ln(\tilde{\mu}^i_{c,t})$	-0.015*	-0.053***	-0.055	-0.026**	-0.037***
	(0.008)	(0.012)	(0.121)	(0.010)	(0.013)
Tech. from OES2016	0.051	-0.823***	-2.517**	-0.335**	-0.288
	(0.176)	(0.173)	(0.955)	(0.161)	(0.190)
Tech. from Mexico (ENOE)	0.029*	-0.028	-0.094	0.014	0.023
	(0.017)	(0.022)	(0.108)	(0.022)	(0.032)
Aggregation of countries	-0.103	-0.952***	-3.140***	-0.429**	-0.529**
	(0.197)	(0.217)	(0.971)	(0.171)	(0.229)
Alternative concordance	0.021	-0.853***	-3.826***	-0.411*	-0.313
	(0.204)	(0.174)	(0.934)	(0.221)	(0.231)
Probit	-0.335	-3.234***		-1.658***	-1.963*
	(0.743)	(0.710)		(0.642)	(1.101)
Dummies	ct	it	ci	c,i	ct,it

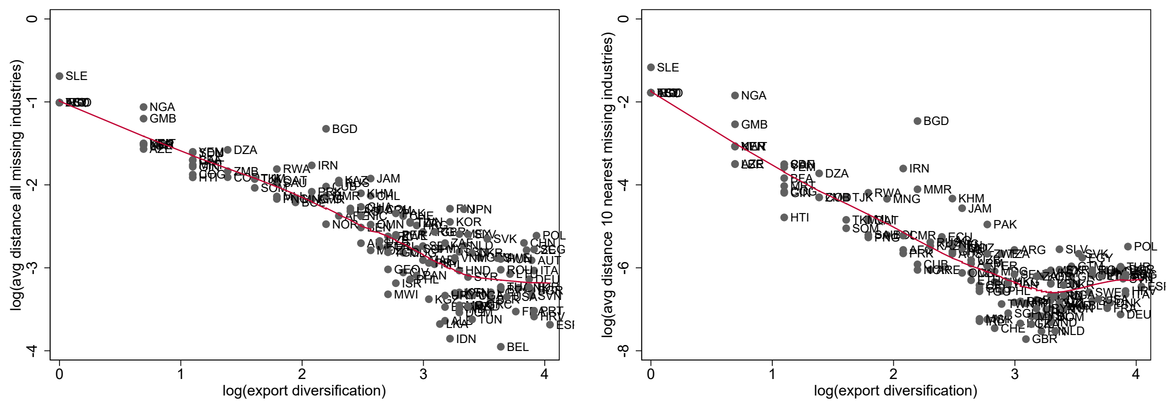
Notes: The dependent variable is survival after entry of a new industry ($y^i_{c,t}$), as in Equation (16). The independent variable is $\tilde{\mu}^i_{c,t}$ (Cobb-Douglas), that is, using the wage-bill share for μ^i_τ. Rows indicate the following robustness checks: *'Benchmark'* is the reference, as in the main text; *'Alternative RCA'* uses 0.5 instead of 1 as the threshold; *'Fixed threshold'* requires a minimum of 10M USD to consider a country an exporter of a commodity; *'$distance^i_{c,t} = \ln(\tilde{\mu}^i_{c,t})$'* uses the logarithm of our main control; *'Tech from OES2016'* uses BLS's occupational employment statistics from 2016 to compute μ^i_τ; *'Tech from Mexico (ENOE)'* uses a dataset from Mexico (Encuesta Nacional de Ocupación y Empleo) to compute μ^i_τ; *'Aggregation of countries'* drops the small countries in the first column of Table S3; *'Alternative concordance'* uses an alternative method to convert HS to NAICS, as described in Supplementary Material S2.5.1; *'Probit'* uses a probit estimator. Note that there are too many dummies in model (3) for this estimator. Country-level cluster robust standard errors in parentheses. Significance: $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$.

Only in one case we find a positive and (weakly) significant coefficient. In all other cases the coefficient is either insignificant or negative and significant. Notably that is the case for column (5), our preferred specification.

S2.3 Diversification and Distance from Missing Industries

Figure S2 shows that the negative relationship between a country's export diversification and its distance from missing industries in Figure 6 is robust to considering only industries with non-zero distance.

Figure S2: Robustness: Export diversification and average distance from missing industries



Notes: The horizontal axis is a country's log export diversification using our threshold for a country to be active in an industry as discussed in Section 2 and data from years 2012-2016. The vertical axis is a country's log average distance from its missing industries using wage-bill shares (Cobb-Douglas). The left panel takes the average over all missing industries with non-zero distance. The right panel the average over the 10-nearest missing industries with non-zero distance. The trend lines depict a LOWESS smooth.

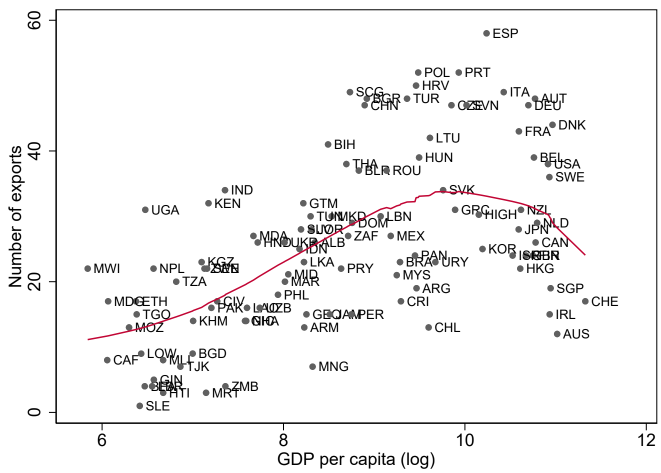
S2.4 Income and Export Diversification

A key premise of our work is that economic diversification tends to increase with income. This observation is well established in the literature and has been documented by Imbs and Wacziarg (2003); Cadot et al. (2010) and others. In this appendix, we document that the same is also true when using the data underlying our empirical analysis in the main text.

In this paper, we employ a modified version of UN Comtrade data where we match export commodities to 88 4-digit North American Industry Classification System (NAICS) industries—see Section 2 and Appendices C and S2.5 for details on the data. In Figure S3, we use this data and plot a country's export diversification against its GDP per capita, taken from the world development indicators (WDI). Export diversification is defined as the number of industries with $RCA > 1$, in line with our definition of presence in the main text. We clearly see a positive association in our data between export diversification and income. We also observe a small decline at the very top (above $\sim \$35,000$). This is, however, not enough to reverse the trend, and no highly diversified country has low levels of income according to our data.[49]

[49]Moreover, the 'reconcentration' at the top appears to be driven by few, relatively small outliers.

Figure S3: Scatter plot of number of exports (div_c) against log GDP per capita



Notes: The indicators are calculated using our baseline sample and averaging data over the period 2012-2016. Countries excessively relying on oil exports (RCA over 5) are excluded from the sample. The trend line depicts a LOWESS smooth.

S2.5 Further Details on the Data

S2.5.1 Concordance Between NAICS and HS

To back out the revealed number of occupations by country from the data, we match international trade data (as a proxy for production) to 88 NAICS industries, for which we observe occupational inputs from the BLS. Pierce and Schott (2009) provide a concordance table that can uniquely match 10-digit HS codes to 6-digit NAICS. In our case, however, we need to map 6-digit HS to 4-digit NAICS. This results in a many-to-many mapping, where in several instances 6-digit HS codes are linked to more than one 4-digit NAICS industry. To resolve these cases, we use two different methods that allow us to assign each 6-digit HS code to a single 4-digit NAICS industry, as we now explain. Our results are robust to the choice of the concordance—see row 'Alternative concordance' of Tables S1 and S2.

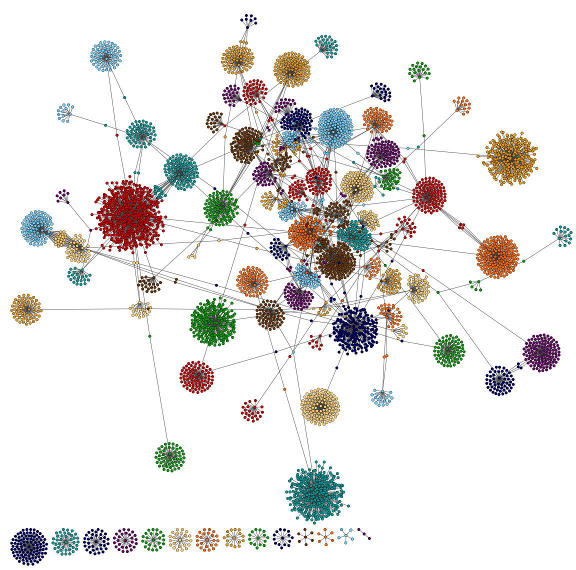
Method 1 - Most frequent NAICS (main method)

For this method, we exploit the mapping of 10-digit HS codes to 6-digit NAICS (Pierce and Schott, 2009). Specifically, for each link from a 6-digit HS code to a 4-digit NAICS code, we count the number of links from underlying 10-digit HS codes to the 4-digit NAICS code. We then assign 6-digit HS codes with ambiguous mapping to the 4-digit NAICS industry that has the highest number of incoming links from the underlying 10-digit HS codes.

Method 2 - Clustering

An alternative procedure is tested for robustness. Based on the observation that concordance tables can be seen as a network where the classifications are nodes and the mappings are edges, it is shown in Diodato (2018) that the clustering of the concordance network can be exploited to resolve multiple mappings. By passing the concordance table into a label propagation algorithm, we detect exactly 88 communities, one for each NAICS industry. We then link the 6-digit HS code to the NAICS industry in the same community. Figure S4 provides a visualization of the network.

Figure S4: Concordance between NAICS and HS



Notes: The figure shows the clustering alogithm's assignment of 6-digit HS codes to a single 4-digit NAICS code, using a different color for each cluster (NAICS code).

S2.5.2 Country Selection and Border Redefinition

Our period of analysis—from 1992 to 2016—has the advantage of being relatively stable in terms of border re-definitions. Notable exceptions exists, from real political changes (such as the separation of Montenegro and South Sudan) to changes in accounting (such as the trade of San Marino and the Vatican). To minimize errors, we aggregate them for the whole period under analysis. Table S3 lists these countries and how they are re-assigned. For robustness, we check if our results hold when dropping these countries outright. The rows 'Aggregation of countries' in Tables S1 and S2 show that this is the case.

Table S3: Re-assignment in the classification of countries

Country	Merged with
Luxembourg	Belgium
Réunion	France
Saint Barthélemy	France
Martinique	France
Guadeloupe	France
French Guiana	France
Timor-Leste	Indonesia
Holy See (Vatican City State)	Italy
San Marino	Italy
Netherlands Antilles	Netherlands
Bonaire, Saint Eustatius and Saba	Netherlands
Curaçao	Netherlands
Saint Maarten (Dutch part)	Netherlands
Bouvet Island	Norway
Montenegro	Serbia and Montenegro
Serbia	Serbia and Montenegro
Lesotho	South Africa
Swaziland	South Africa
Botswana	South Africa
Namibia	South Africa
South Sudan	Sudan
United States Minor Outlying Islands	USA
Guam	USA

Notes: All countries on the left column have had (at some point in time during the period of analysis) their trade recorded together with the corresponding country on the right column. Merging their records for the whole 1992-2016 period removes bumps in the data.

References

Cadot, O., Carrère, C., and Strauss-Kahn, V. (2010). Export diversification: What's behind the hump? *Review of Economics and Statistics*, 93(2):590–605.

Diodato, D. (2018). A network-based method to harmonize data classifications. Papers in Evolutionary Economic Geography 1843.

Imbs, J. and Wacziarg, R. (2003). Stages of diversification. *American Economic Review*, 93(1):63–86.

Kremer, M. (1993). The O-ring theory of economic development. *Quarterly Journal of Economics*, 108(3):551–575.

Pierce, J. R. and Schott, P. K. (2009). A concordance between ten-digit U.S. harmonized system codes and SIC/NAICS product classes and industries. Working Paper 15548, NBER.

Schetter, U. (2020). Quality differentiation, comparative advantage, and international specialization across products. Research Fellow and Graduate Student Working Paper 126, CID at Harvard University. http://dx.doi.org/10.2139/ssrn.3091581.