TI 2014-084/VI
Tinbergen Institute Discussion Paper

Does Access to Foreign Markets shape Internal Migration? Evidence from Brazil

Laura Hering[1]
Rodrigo Paillacar[2]

[1] *Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute, the Netherlands;*
[2] *University of Cergy-Pontoise, France.*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: http://www.dsf.nl/

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

Does Access to Foreign Markets Shape Internal Migration? Evidence from Brazil

Laura Hering[*] Rodrigo Paillacar[†]

July 7, 2014

Abstract

This paper investigates how internal migration is affected by Brazil's increased integration into the world economy. We analyze the impact of regional differences in access to foreign demand on sector-specific bilateral migration rates between the Brazilian states for the years 1995 to 2003. Using international trade data, we compute a foreign market access indicator at the sectoral level, which is exogenous to domestic migration. A higher foreign market access is associated with a higher local labor demand and attracts workers via two potential channels: higher wages and new job opportunities. Our results show that both channels play a significant role in internal migration. Further, we find a heterogeneous impact across industries, according to their comparative advantage on the world market. However, the impact of market access is robust only for low-educated workers. This finding is consistent with the fact that Brazil is exporting mainly goods that are intensive in unskilled labor.

Keywords: Regional migration, international trade, market access, Brazil.

JEL classification: F16, F66, R12, R23.

[*] Erasmus University Rotterdam & Tinbergen Institute, the Netherlands.
[†] University of Cergy-Pontoise, France.

1 Introduction

A considerable amount of literature provides evidence that a country generally benefits from opening up to international trade. However, within the country, these benefits are often unevenly distributed. This can cause a rise in regional wage disparities, both across and within industries, which may lead to changes in the spatial distribution of the domestic economic activity. Today, a growing number of studies is exploring the consequences of globalization on the sectoral and spatial adjustment of the domestic labor force within countries.[1]

In this paper, we investigate how internal migration is affected by Brazil's increased integration into the world economy. More specifically, we analyze the impact of changes in foreign demand for Brazilian goods on sector-specific bilateral migration rates between the 27 Brazilian states for the years 1995 to 2003.

In order to identify the effect of international trade on the local labor market in a specific sector, we use a region-sector specific measure of foreign demand. We compute a theory-based measure of access to foreign markets, which is derived from a standard gravity equation that can be obtained from various trade models. The location of the region with respect to its potential trading partners plays a key role in determining a region's market access. Firms located in regions closer to large consumer markets have a higher market access due to lower trade costs, thereby giving them a competitive advantage in these markets. An increase in a region's market access therefore reflects a higher demand for its products and consequently a higher labor demand.

The main contribution of this paper is to show that an increase in a region's access to

[1] See the World Bank Development Report (2009) for a global picture on the impact of international trade on spatial disparities within countries. More recently, Topalova (2010) shows how the effect of India's trade liberalization on poverty reduction varies across districts and Helpman et al. (2012) study the impact of international trade on inequality within sectors for the case of Brazil. For an overview on the impact of trade liberalization on inequality in developing countries see Goldberg and Pavcnik (2007).

foreign markets attracts migrants via two channels: i) an indirect effect via an increase in the local wage premium and ii) a direct effect resulting from the creation of new job opportunities.

The positive effect of foreign market access on wages is already well documented for various countries, including Brazil (Fally et al., 2010).[2] In this paper we focus on the second channel, which captures the impact of market access on migration beyond its effect via a change in local wages.

Higher market access is expected to have also a direct effect on migration essentially due to a higher number of vacancies, which increases the probability of employment. Moreover, the type of jobs created as a result of an increased foreign demand can be considered to be of better quality, thus representing an additional incentive for migration. In Brazil, as in many emerging countries, firms in the export industry are preferred employers. They are likely to offer more long-term employments, propose a steeper wage gradient and better working conditions.[3] An increase in the market access variable can thus also capture long-term considerations in the migration decision. These aspects are typically excluded when migration is modeled as depending only on spot wages, which themselves cannot capture the workers' wage profile or non-pecuniary aspects linked to the job (Aguayo-Tellez et al., 2010).

Two features make Brazil a particular interesting case to study. First, the country experienced a period of important trade liberalization in the early 1990s, which was accompanied by a significant change in its internal economic geography. Volpe-Martincus

[2] The impact of market access on wages is by now well studied empirically. See, among others, Hanson (2005) for the US, Head and Mayer (2006) for Europe, and Hering and Poncet (2010) for China. The theoretical link is modeled explicitly in the so-called "New Economic Geography wage equation" (Fujita et al., 1999), but Head and Mayer (2011) point out that such wage equations can be established in numerous trade models.

[3] While many studies have documented that exporters pay higher wages (see Wagner (2012) for an overview), Krishna et al. (2014) have found that in Brazil the exporter wage premium can be fully explained by a better matching of workers and jobs in exporting firms with respect to non-exporters.

(2010) highlights that the change in trade policy has especially favored the creation of firms located in Brazilian states close to Argentina, Brazil's second most important trading partner. Second, Brazil exhibits exceptionally high levels of internal migration: in 1999, 40% of all Brazilians lived in a state other than the state in which they were born in (Fiess and Verner, 2003).[4]

Our empirical specification is derived from an aggregate discrete choice model, where bilateral migration is modeled as depending on utility differentials in a gravity framework.

We construct a sector-specific foreign market access indicator with the estimates of a gravity trade equation, following the methodology proposed by Redding and Venables (2004). The main advantage of this measure is that it identifies the effect of foreign demand on the local labor market. Note that a positive shock to foreign market access does not necessarily mean that only jobs in exporting firms will be created. Due to spillovers or an increase in connected activities (e.g. outsourced tasks), the increase in demand for exported goods can also lead to a higher labor demand in non-exporting local firms in the same sector. These employment opportunities are also captured by our indicator.

A further advantage is that our theory-based market access measure is by construction exogenous to domestic factors, such as local labor market regulations, that may affect wages and migration. Therefore, it also eliminates the possibility of reverse causality between internal migration and international market access.

Performing the analysis of bilateral migration at the sectoral level is motivated by some recent studies on Brazil's labor market, which present evidence for a very low sectoral mobility of Brazilian workers (Menezes-Filho and Muendler, 2011; Muendler, 2008).

[4] For comparison: Dahl (2002) finds that in 1999 in the US, a country known for its high labor mobility, 30% of all male workers no longer live in the state in which they were born.

Therefore, in this paper we focus on labor migration that takes place within sectors.[5]

The sectoral approach has two important advantages, which we exploit in our identification strategy. First, in contrast to our sectoral measure, an aggregated market access variable would be potentially correlated with the evolution of other unobserved migration determinants that vary over time and across states (i.e. amenities, price levels, institutional quality). Constructing migration rates and market access by sector allows us to include year-location fixed effects, which control for these unobserved location characteristics. Second, an analysis at the sectoral level allows us to study the heterogeneous effect of market access across industries.

Our results show that regional differences in access to international markets indeed affect internal migration patterns, suggesting that workers migrate from states with low market access to states with high market access. We provide evidence that foreign demand impacts migration also directly and not only by means of an increased wage level. These findings suggest that new job opportunities created by higher foreign demand are important location determinants.[6]

We have access to several rounds of Brazilian household surveys (PNAD). This enables us to construct bilateral migration rates not only by sector but also by educational attainment, which altogether allows us to detect heterogeneous responses to changes in market access across sectors and across educational levels. When looking at differences across sectors, we find that access to international markets has a robust impact in eight out of eleven sectors. Our results indicate that the effect of market access is generally stronger, the higher the industry's comparative advantage is on the world market.

[5] In the empirical analysis, we provide some robustness checks accounting for potential sectoral reallocation.

[6] Due to the lack of more detailed data on local labor market characteristics, we cannot identify the main driver of the observed direct effect (i.e. higher number of jobs, better career opportunities or better working conditions).

Moreover, we find that the impact of market access on sectoral migration rates is robust only for low-educated workers. This could be explained by Brazil's relative abundance of low-skilled labor. A higher market access represents a stronger increase in demand for goods intensive in low-skilled labor, in which Brazil has a comparative advantage on the world market (Muriel and Terra, 2009). Thus, these workers are more likely to be affected by a change in the foreign demand.

Finally, to aid interpretation of our estimated coefficients, we simulate the implied changes in immigration shares for each Brazilian state in the presence of specific positive shocks to foreign demand. This illustrates well how the the spatial structure of the domestic economy is affected by international trade.

Although several studies explore the link between trade and migration, they have mostly focused on international migration patterns (cf., for example, Ortega and Peri, 2013, and Letouzé et al., 2009). Yet, internal migration flows have a far greater magnitude than international flows, and hence may modify a country's development path much more sensibly. This is of particular relevance in fast urbanizing developing countries like Brazil, as highlighted by the World Bank (2009).

Closest to our work is the paper by Aguayo-Tellez et al. (2010), which also applies to Brazil. These authors analyze how the presence of exporters and multinational enterprises in a given state affects the individual migration decision. They find that workers in formal employments are attracted by a higher concentration of foreign owned establishments. We differ from that paper in several aspects. Most importantly, we focus only on employment opportunities that are created by a change in foreign demand. However, as explained above, these new vacancies can also be in non-exporting and domestically owned firms.[7]

[7] In our empirical analysis, the presence of exporters and foreign owned firms is controlled for via location-year fixed effects.

Further, our analysis also includes informal workers, who account for at least 38% of the Brazilian workforce (Henley et al., 2009).

A few papers have studied the role of imports in the location choice of individuals and can be considered as complementary to our work. Crozet (2004) uses a New Economic Geography framework to explain migration between European regions as a function of real wage differentials, where each region's price level depends on the prices of imports. Kovak (2011; 2013) studies the effect of import competition on internal migration patterns. He shows in a specific-factor framework how regions within Brazil are differently affected by trade liberalization. He finds that regions specialized in industries experiencing larger tariff cuts see their wages decrease, which in turn triggers outmigration. In the same spirit, Autor et al. (2013) show how import competition from China affects local labor markets in the United States. They find that commuting zones with stronger import competition experience a higher reduction in manufacturing employment. However, their setting requires internal migration in reaction to trade shocks being negligible.[8]

The remainder of the paper proceeds as follows: Section 2 outlines the estimation strategy for the migration equation; Section 3 formally derives the market access indicator and presents its construction. Section 4 details the migration data and how we obtain our wage variable. Section 5 presents the main results. In Section 6, we present the results by educational attainment and simulate for each state the migration response to a shock in the foreign demand. Section 7 concludes.

[8] Note also that their proxy of trade exposure is only region-time specific. Since we exploit the sectoral dimension and control for location-time fixed effects, we automatically account for this measure.

2 Empirical methodology

The empirical specification of our migration equation is based on an additive random utility model.[9]

Every individual k from location i maximizes the indirect utility V_{kij} across all possible destinations j. In a general utility differential approach, the individual location choice M_{kij} can then be written as:

$$M_{kij} = 1 \quad if\ and\ only\ if \quad V_{kij} = max\ (V_{ki1}, \ldots, V_{kiJ}),$$

$$= 0 \quad otherwise.$$

The indirect utility V_{kij} can be decomposed as follows:

$$V_{kij} = X_{ij}\beta + \xi_{ij} + e_{kij} \tag{1}$$

where X_{ij} are the characteristics of location j. The subscript i is included, as characteristics of j can vary across original locations i (e.g. bilateral distance). β is a vector of marginal utilities and ξ_{ij} represents unobserved location characteristics. The idiosyncratic error term e_{kij} is included to allow individuals from the same origin to choose different locations. We make the standard assumption that this error term follows an i.i.d. Type I extreme value distribution.

Given that individuals select the location that maximizes their utility, the probability that an individual from i will choose destination j is defined by

$$Pr(V_{kij} > V_{kim}) \quad \forall j \neq m \tag{2}$$

[9] This model choice is standard in the recent migration literature and is used for example in Grogger and Hanson (2011) and Kovak (2011). For a detailed description on the derivation of the empirical specification see Bertoli and Fernández-Huertas Moraga (2013).

Replacing the indirect utilities by their definitions of Equation 1 and rearranging terms, the probability that individual k will move from i to j is given by:

$$Pr(e_{kij} - e_{kim} > X_{im}\beta - X_{ij}\beta + \xi_{im} - \xi_{ij}) \ \ \forall j \neq m \tag{3}$$

McFadden (1974) shows that under the assumption of an i.i.d. extreme value distribution of the individual error term, migration probabilities can be expressed as

$$Pr(M_{kij} = 1) = \frac{exp(X_{ij}\beta + \xi_{ij})}{\Sigma_{j=1}^{J} exp(X_{ij}\beta + \xi_{ij})} = s_{ij} \tag{4}$$

Following Berry (1994), this individual migration probability can be interpreted as the share of individuals from i migrating to j, s_{ij}. Similarly, the share of stayers of region i, s_{ii}, can be written as

$$Pr(M_{kii} = 1) = \frac{exp(X_{ii}\beta + \xi_{ii})}{\Sigma_{j=1}^{J} exp(X_{ij}\beta + \xi_{ij})} = s_{ii} \tag{5}$$

Dividing Equation 4 by Equation 5 and taking the log yields

$$ln(\frac{s_{ij}}{s_{ii}}) = ln\left(\frac{exp(X_{ij}\beta + \xi_{ij})}{exp(X_{ii}\beta + \xi_{ii})}\right) = \beta(X_{ij} - X_{ii}) + \xi_{ij} - \xi_{ii} \tag{6}$$

We now have an aggregate discrete choice model that accounts for unobserved location characteristics ξ and whose parameters can be estimated using conventional linear estimation techniques.

To obtain our empirical specification, we add the time dimension t and the sectoral dimension s and replace the vector X with our location-sector specific variables of interest. Here we make the implicit assumption that workers do not switch sectors, and thus their migration decision depends only on state characteristics (e.g. price level) or the characteristics of their own sector (e.g. sectoral market access). This gives us our first benchmark specification:

$$\ln m_{ijst} = \ln \frac{s_{ijst}}{s_{iist}} = \alpha + \beta_1 \Delta MA_{ijs\tilde{t}} + \beta_2 \Delta w_{ijs\tilde{t}} + FE_{ij} + FE_{st} + FE_{it} + FE_{jt} + \varepsilon_{ijst} \quad (7)$$

m_{ijst} is the observed migration rate between state i and j for sector s in the household survey of year t. It is simply defined as the number of migrants going from i to j divided by the number of stayers. Individuals are considered as migrants when they declare having lived five years ago $(t-5)$ in a different state than their current state of residence. Since we do not know the exact moment of migration, all independent variables are constructed as means over the years $t-4$ to $t-1$. This is indicated by the index $\tilde{t}$.[10]

Our main variable of interest is the market access gap between states i and j for sector s, $\Delta MA_{ijs\tilde{t}} = ln\frac{MA_{js\tilde{t}}}{MA_{is\tilde{t}}}$. An increase in this variable makes state j relatively more attractive, either because of i) a higher wage level or ii) new job opportunities (more or better jobs). We can isolate the second channel by including the wage gap, $\Delta w_{ijs\tilde{t}}$, in our benchmark specification. Adding the wage variable has an additional important advantage: it also captures other sector and time varying characteristics of the local labor market that we cannot observe, but which are potentially correlated with foreign market access (e.g. sector-specific productivity differentials).

We also provide results without the wage gap to estimate the joint impact of both channels. However, these results are more likely to suffer from an omitted variable bias, since they do not control for changes in other sector-specific determinants of labor demand.

A lower number of available jobs is typically corresponding also to a higher unem-

[10] Our benchmark results hold also when specifying our independent variables as four-year lags instead of the mean over the previous four years.

ployment rate. But a higher unemployment rate can also reflect limitations on the labor supply side or a mismatch on the local labor market between vacancies and job seekers. While in some specifications we explicitly include regional differences in unemployment rates, our benchmark estimation includes FE_{it} and FE_{jt}, which correspond to origin-year and destination-year dummies. These account for time-varying differences across states, including the unemployment rate, amenities or price levels, which are also considered to be important determinants of migration.

Bilateral fixed effects FE_{ij} take into account time-invariant specificities concerning migration between two particular states (e.g. moving costs, migration networks). FE_{st} represents sector-year fixed effects.

By definition, ε_{ijst} is a i.i.d. bilateral error term. However, using Equation 6 it can be shown that all ε_{ijst} from the same origin i depend on the same ξ_{ii}. This leads to a non-zero covariance of ε_{ijst} for observations with the same origin i in year t. In all our regressions, we therefore cluster our standard errors by the state of origin-year level.[11]

A last feature of our model we have to take into account is the underlying assumption of the independence of irrelevant alternatives (IIA). The probability of choosing one state relative to the probability of choosing another may be dependent on the characteristics of a third state. If two or more destinations are perceived as close substitutes by potential migrants, this leads to the violation of the assumption of the IIA. We control for the potential correlation across alternative destinations through the inclusion of origin-time fixed effects as proposed by Ortega and Peri (2013). The additional bilateral and destination-time fixed effects further reduce the risk of obtaining biased coefficients due to a violation of the IIA.[12]

[11] The choice of clusters does not affect our results. We specify different levels of clusters, but significance levels stay highly similar.

[12] Bertoli and Fernández-Huertas Moraga (2013) caution that problems may remain also in this case

3 Market access: derivation and construction

3.1 Theoretical derivation of market access

In this subsection, we provide the formal definition of market access and how it can be derived from a standard gravity model of trade. Although initially derived from a trade model of monopolistic competition, this indicator is supported by various different micro foundations of the trade gravity equation.[13]

According to structural gravity models, exports EX_{ijs} in sector s from region i to partner j can be written as

$$EX_{ijs} = \phi_{ijs} S_{is} M_{js} = \phi_{ijs} \underbrace{\frac{Y_{is}}{\Pi_{is}}}_{S_{is}} \underbrace{\frac{E_{js}}{P_{js}}}_{M_{js}} \tag{8}$$

with $0 \leq \phi_{ijs} \leq 1$. This equation decomposes exports into three components: The term ϕ_{ijs} reflects the accessibility of market j for the exporters from location i in sector s. A ϕ_{ijs} of 1 indicates free trade and $\phi_{ijs} = 0$ refers to prohibitively high trade costs and thus zero exports.

The terms S_{is} and M_{js} are often referred to in the literature as the supply and market capacity. They capture all the considerations that make exporter i a competitive exporter and partner j an attractive destination in sector s. More precisely, the supply capacity depends on the total output $Y_{is} = \sum_j EX_{ijs}$ of sector s in location i, as well as the local

and suggest to test for the presence of cross-sectional dependence in the residuals. Unfortunately, our panel is too short to allow for such tests. However, Hausman and McFadden (1984) note that if the IIA is satisfied, then the estimated regression coefficients should be stable across all 27 destination sets. We thus re-estimate our model 27 times, each time dropping one of the destinations (following Grogger and Hanson, 2011). Coefficients of our market access and wage variables stay similar across samples, suggesting that the IIA property is not violated here.

[13] This subsection borrows from the presentation of the general framework in Head and Mayer (2013). They show how market access can be derived also in other market structures, notably in a setting with perfect competition and technology differences (Eaton and Kortum, 2002), or in trade models accounting for firm heterogeneity (Chaney, 2008).

firms' price competitiveness, Π_{is}. The market capacity of j in sector s depends on location j's total expenditure on goods from sector s, $E_{js} = \sum_i EX_{ijs}$, and the prevailing price index in sector s on market j, P_{js}.

The terms Π_{is} and P_{js} are the so-called outward and inward "multilateral resistance terms" (Anderson and van Wincoop, 2003). These terms take into account that bilateral trade relationships are affected by competition from third countries.

Given Equation 8, region i's relative access to every individual market j for sector s is defined by $\frac{E_{js}\phi_{ijs}}{P_{js}}$. Region i's total market access in sector s can be obtained by summing over all destinations j:

$$MA_{is} = \sum_j \frac{E_{js}\phi_{ijs}}{P_{js}} = \sum_j \phi_{ijs} M_{js} \tag{9}$$

MA_{is} measures the overall ease for firms in location i to access all domestic and foreign markets j in sector s. This indicator represents an expenditure-weighted average of relative access, as it weights the market capacity of each potential destination j by their accessibility from region i.

By summing only over foreign countries, we obtain an international market access indicator, which solely captures the demand for goods from location i coming from abroad. This allows us to identify the effect of international markets on the domestic economy. The market access indicator differs from total exports since it excludes the local supply capacity. This way, the market access measure is exogenous to local factors that affect the overall export capacity of the region, such as the number of exporting firms in sector s or locally available technology.

A further advantage of focusing on foreign market access is that it eliminates possible reverse causality problems that arise when immigrants raise local consumption and hence local demand: a local shock inducing the arrival of additional migrants may increase consumption in the host region and thus domestic market access, but does not affect the access to foreign markets.

3.2 Market Access Calculation

We estimate the market access indicator presented in Equation 9 via a gravity trade regression, following Redding and Venables (2004). This methodology is rarely applied in regional studies because of data limitations: bilateral trade flows are often unavailable at the subnational level, particularly for developing countries. Brazil is a fortunate exception since it provides information on international trade flows at the sectoral level for each of its 27 states.

Our trade data set covers the years 1991 to 2002 and 11 sectors.[14] It contains international trade flows between the 27 Brazilian states and 170 partner countries and flows among the 170 foreign countries.

The empirical specification of the trade equation follows from Equation 8. After taking the logs, we obtain

$$\ln EX_{ijs} = \ln \phi_{ijs} + \ln S_{is} + \ln M_{js} \tag{10}$$

For the calculation of a sector-state specific market access variable that varies over time, we estimate Equation 10 separately for every sector-year pair.

In the regressions, sector-specific market capacity (M_{js}) and supply capacity (S_{is}) of every trading partner are captured by sector-importer (FM_{js}) and sector-exporter (FX_{is}) fixed effects. ϕ_{ijs} can be specified using different measures of trade costs. Specifically,

[14] For details on sectoral classification and data sources for the variables used in this section see Appendix B-1 and B-2.

we consider bilateral distance (d_{ij})[15], whether partners share a common border (B_{ij}), the presence of a free trade agreement between the two trading partners (RTA_{ij}) and whether the two are member of the WTO or its predecessor GATT (WTO_{ij}).

Since we estimate the trade equation separately for every sector-year pair, we can drop the subscript s. Our empirical specification of the trade equation can then be written as

$$\ln EX_{ij} = \delta \ln d_{ij} + \lambda_1 B_{ij} + \lambda_2 RTA_{ij} + \lambda_3 WTO_{ij} + FX_i + FM_j + \nu_{ij} \tag{11}$$

where ν_{ij} is a random bilateral error term.

In total, we run 132 regressions (12 years $\times$ 11 sectors). Given that all coefficients and fixed effects are allowed to vary over time and across sectors, this enables us to build a time-varying market access indicator specific for each state-sector combination.

Market access for state i in sector s in year t is built by weighting each predicted market capacity, $\widehat{M_{jst}}$, by the estimates of the corresponding bilateral trade costs, $\widehat{\phi_{ijst}}$. These weighted market capacities are then summed up to one single indicator per state-sector pair:

$$\begin{aligned} MA_{ist} &= \textstyle\sum_j \widehat{\phi_{ijst}} \widehat{M_{jst}} \\ &= \textstyle\sum_j^R \exp(\widehat{\delta_{st}} \ln d_{ij} + \widehat{\lambda_{1st}} B_{ij} + \widehat{\lambda_{2st}} RTA_{ijt} + \widehat{\lambda_{3st}} WTO_{ijt} + \widehat{FM_{jst}}) \end{aligned} \tag{12}$$

We sum over R countries, where R includes only foreign countries and not the Brazilian states. This way, our market access measure exclusively captures the $foreign$ demand addressed to each Brazilian state.[16]

[15] The distance between a Brazilian state and a partner country is defined as the sum of two components: the internal distance between the state and the closest sea harbor and the external distance between the harbor and the foreign country. See Appendix B-2.1 for a detailed description of the construction of bilateral distances.

[16] To be consistent across sectors and years, each MA_{is} is constructed using the estimated market capacities and trade costs of always the exact same 92 countries. These are the countries that import goods from all sectors in all years and thus provides us for all sector-year combinations with the necessary estimates for trade costs and importer fixed effects.

For robustness, we compute various different market access indicators. First, we estimate Equation 11 using different specifications for ϕ_{ij}.[17]

Second, the indicator MA_H addresses the concern of zero-value trade flows. We follow the methodology proposed by Helpman et al. (2008) and control for zeros by adopting a Heckman two-step estimator, using a bilateral "doing business" indicator as selection variable in the first step (see Appendix B-2).

Finally, since not only foreign, but also domestic demand should matter for the migration decision, we construct a total market access indicator (MA_T) that includes also a domestic dimension. As a proxy for the spatial distribution of domestic demand, we include here the trade cost weighted importer fixed effects of the Brazilian states.

3.3 Results of the trade equation and market access description

Table A-1 in Appendix A summarizes by industry the coefficients obtained from the trade regressions (Equation 11). Coefficients on the trade cost variables have the expected sign and magnitudes are in line with the literature (cf. Head and Mayer, 2013). However, there are some important differences across sectors, in particular in the distance coefficient.

The last column summarizes the time varying importer fixed effect, representing the sector-specific market capacity of each destination country. When comparing the average market capacity by importer, we see substantial variations, with the US and Japan being the two countries with the highest average market capacity. Interestingly, we see for Brazil's most important neighbor, Argentina, a substantial increase in average market capacity during the early 1990s, followed by a big drop after 1999, which coincides with

[17] In Table 2 two additional market access measures are employed: Indicator MA_1 uses the geodesic distance as bilateral distance between the Brazilian states and the foreign trading partners. MA_2 uses the same definition of the bilateral distance as our preferred specification, but estimates the trade equation using only common border and distance as variables for bilateral trade costs.

the Argentinian economic crisis.

The differences in the estimates of trade costs and market capacities across sectors and years result in a different market access ranking of states across sectors.

Figure A-1 displays a map of the Brazilian states.[18] Figure A-2 shows the spatial distribution of market access across states (reporting averages over years and sectors). Rio Grande do Sul in the South, bordering Uruguay and Argentina, has the highest average market access over our sample period, followed by its neighbor Santa Catarina, which is also sharing a border with Argentina. Due to its short internal distance to a harbor and its proximity to the important markets in North America and Europe, Amapá on the Northern coast comes third. The state with the lowest market access average is Mato Grosso in the Center-West, which has no direct access to the sea.

Tables A-2 and A-3 report summary statistics and the correlation between our main variables. The correlation of our preferred indicator $\Delta MA_{ijs\tilde{t}}$ with the other market access variables is overall very high.

4 Household survey data

Our main data set is the yearly household survey *Pesquisa Nacional por Amostra de Domicilios* (PNAD), which covers between 310,000 and 390,000 individuals each year. The PNAD does not follow individuals, but interviews a different random and representative sample of residents each year. The data is collected and published by the Brazilian

[18] The Brazilian states are grouped in five macro-regions. This classification is based on the structural and economic development of the different states, regrouping states with similar characteristics. The North is sparsely populated and largely inaccessible. The Northeast is the poorest macro-region of Brazil with the lowest life expectancy and wages and the highest proportion of low educated persons. The Center-West combines a diverse set of characteristics, mixing poor rural areas, dense forests, and the federal capital city of Brasilia, where income and education levels are high. The Southeast and the South are the most economically developed regions of Brazil.

Institute of Geography and Statistics (IBGE). In this study, we use the PNAD for the years 1992 to 2003 (with data missing for 1994 and 2000).[19]

In the following, we present the construction of the bilateral migration rates and of the wage gap using the PNAD.

4.1 Migration rates

We identify an individual as a migrant, when the answer given to the question "In which Brazilian state did you live five years ago?" differs from the actual state of residence.

Our sample is limited to individuals who declare having a job in a tradeable sector, earning a positive wage, having lived in Brazil 5 years ago and being between 20 and 65 at the time of the interview.[20]

We distinguish 11 tradeable sectors that can be matched with the trade data and construct bilateral migration rates separately for each sector.[21] We do not have any information about the individual's work five years ago. Nevertheless, as argued above, we can make the reasonable assumption that individuals already worked in the same sector as in the year of the survey. Bilateral migration rates are then defined as the number of migrants from state i to j over the number of workers that stayed in state i and declare working in sector s at the time of the interview. In Section 6, we construct sectoral migration rates separately by educational attainment. The workers are treated as highly educated if they attended high school for at least one year; otherwise they are regarded as low educated. In our empirical analysis, we exclude migration rates that are constructed

[19] In 1994 the PNAD was not conducted because of a strike. 1991 and 2000 were years of the population census. We do not use the PNAD data before 1991, because it does not provide the necessary detailed industry classification.

[20] Unreported results (available upon request) using a sample limited to household heads are overall very similar.

[21] See Appendix B-1 for details on the industrial classification.

with less than six observations.[22]

Despite the presence of a relatively high number of zero migration flows among the states, the PNAD is considered to be representative of overall migration rates and thus adequate for studying migration patterns within Brazil (Fiess and Verner, 2003; Cunha, 2002). Nevertheless, in robustness checks, we will address the problem of unobserved flows by running Poisson-Maximum-Likelihood estimations including zero-flows.

In our final data set, close to 3% of the individuals have moved states at least once within five years prior to the interview. Even though most of the migrants are low qualified in absolute terms, the highly educated individuals are the more mobile throughout the years (2.75% versus 3.53%).

Table A-4 in Appendix A compares interstate mobility across sectors. Whereas less than 3% of the workers in basic metals, machinery, textile and agriculture migrated within the last five years, this percentage is above 4% in the wood industry.

4.2 Sector-state specific wages

Our key control variable is the sector-specific wage gap, $\Delta w_{ijs\tilde{t}}$. This variable accounts for most sector-state specific characteristics of the Brazilian labor market (such as sectoral and regional variations in employment regulations and labor productivity). Moreover, it controls for the indirect impact of market access on migration.

However, for the construction of the wage gap, we do not want to use the observed wage level. The personal characteristics (e.g. education, age) that drive the location choice are also major wage determinants. Thus, the observed wage level in a region depends on the composition of the local labor force, including the immigrants. We treat

[22] Results are robust when maintaining all observed flows and when omitting the top five and bottom five percent of migration rates (results available upon request).

this issue by correcting for self-selected migration, following the methodology developed by Dahl (2002).[23]

$\Delta w_{ijs\tilde{t}}$ is constructed from estimates of a modified Mincerian wage equation that is run separately for every state-year combination jt.[24] The obtained parameters on the individual characteristics are then used to predict wages that each individual k would potentially earn in each of the 27 states in year t. The effect of sector-specific market access on the wage level is accounted for by sector fixed effects.

The final wage gap for year t is defined as

$$\Delta w_{ijst} = \widehat{w_{ijst}} - \widehat{w_{iist}} \tag{13}$$

where $\widehat{w_{iist}}$ is the average of the predicted wages that all individuals k in sector s, who actually lived in i five years ago would earn in state i in year t. $\widehat{w_{ijst}}$ uses the same set of individuals k and is defined as the average of the predicted wages in year t that all workers in sector s coming from state i would have potentially earned in state j (regardless of whether in year t they actually live in j or not).

This aggregation method keeps the composition of the labor force constant across the states, since the same individuals are used for computing the regional wage at the origin and at the destination. Thus, differences in regional wage levels are only due to variations in the estimated parameters of the wage equations and not to the composition of the labor

[23] This approach has become standard in the recent migration literature. For a most recent study see Bertoli et al. (2013).

[24] We regress individual hourly wages over the standard wage determinants age, age squared, education, gender, ethnic group and sector dummies plus an individual correction term. Wages are deflated by the Brazilian consumer price index. The correction terms are the individuals' migration probabilities as proposed by Dahl (2002). The individual probability of moving from i to j is constructed using only observed personal characteristics (educational attainment, age group, gender, family status and state of origin). By adding a polynomial of these migration probabilities to the wage equations, we get consistent estimates of the coefficients on the wage determinants. Estimation results of the wage equations are available upon request.

force.[25]

In Section 6, we use migration rates that are constructed separately for highly educated and low-educated workers. Here, the wage variable takes different values for the different educational groups e. Δw^e_{ijst} is constructed as in Equation 13, but takes the average of the predicted wages only for the relevant group of workers.

5 Main results

5.1 Sector-specific market access and migration rates

In this section, we display our main results. In Column 1 of Table 1, we start by estimating a standard model of migration with a reduced set of fixed effects. Moreover, we exclude the market access indicator to see how this basic specification performs without our variable of interest in explaining bilateral migration rates.

Next to sector-specific wage gaps, we also take into account the regional differences in unemployment rates, $\Delta u_{ij\widetilde{t}}$, population size, $\Delta pop_{ij\widetilde{t}}$, and homicide rates, $\Delta death_{ij\widetilde{t}}$.[26] Homicide rates are considered as a proxy for crime and security. For both the unemployment gap and the difference in homicide rates, we expect a negative impact. The expected sign of population is more ambiguous. Although there are more jobs available in large states, there are also possible congestion costs. In Column 1, all coefficients have the expected sign and are significant, except for the population variable.[27]

[25] Note that Δw_{ijst} is constructed using predicted wages in levels and not in log, as do Grogger and Hanson (2011). When repeating our main estimations with the wage variables in log, wages are not significant and market access shows a higher coefficient. Overall, this would not affect our general conclusions on market access. However, given the highly significant results for wages in levels, we believe that wages in this form are the relevant variable for the estimation of the location decision of workers in Brazil.

[26] See Appendix B-3 for the sources of the additional control variables and the construction of the unemployment rate.

[27] We do not adjust standard errors for the fact that the market access indicators and wage variable

Table 1: Sectoral market access and bilateral migration

Dep. variable:	ln(migrants$_{ijst}$/stayers$_{iist}$)						
	(1)	(2)	(3)	(4) *benchmark I*	(5)	(6) PPML	(7)
$\Delta MA_{ijs\widetilde{t}}$		0.750^a	0.617^a	0.571^a	0.745^a	0.983^a	0.846^a
		(0.100)	(0.086)	(0.097)	(0.116)	(0.259)	(0.280)
$\Delta w_{ijs\widetilde{t}}$	0.271^a		0.251^a	0.311^a		0.170^a	0.041
	(0.047)		(0.044)	(0.048)		(0.056)	(0.033)
$\Delta u_{ij\widetilde{t}}$	-0.246^a	-0.275^a	-0.262^a				-0.268^a
	(0.078)	(0.076)	(0.077)				(0.066)
$\Delta pop_{ij\widetilde{t}}$	0.026	0.071	-0.031				-0.368
	(0.764)	(0.640)	(0.745)				(0.621)
$\Delta death_{ij\widetilde{t}}$	-0.131^c	-0.130^c	-0.129^c				-0.078
	(0.075)	(0.070)	(0.074)				(0.061)
FE_{ij}	yes	yes	yes	yes	yes	yes	yes
FE_{st}	yes	yes	yes	yes	yes	yes	yes
FE_{it}&FE_{jt}				yes	yes	yes	
FE_{is}&FE_{js}							yes
Observations	4183	4183	4183	4183	4183	13927	4183

Heteroskedasticity-robust standard errors clustered at the state of origin-year level appear in parentheses. ^{a}, ^{b} and ^{c} indicate significance at the 1%, 5% and 10% confidence levels.

Column 2 includes our variable of interest, but not the wage variable. This specification captures the joint effect market access has on migration via the two possible channels: higher wages and more job opportunities. The coefficient of market access has the expected positive sign and is highly significant, whereas the coefficients of the other variables change minimally.

In Column 3, we include all control variables. As expected, the coefficient of market access decreases, when wages are added, but remains highly significant. The observed effect here corresponds to the impact market access has on migration beyond its indirect impact via the wage gap. This direct effect can be interpreted as a consequence of the growth in the number of vacancies or in the number of "high quality" jobs created by a

are themselves estimated. Bootstrapping standard errors is prohibitive given the already considerable computational requirements for the construction of each of these variables. Notably, for each of the five market access indicators described in Section 3.2 we need to estimate 132 trade equations. The wage variable is constructed with the estimated parameters of 243 wage equations.

higher foreign demand. This increases the utility of workers in this state and thus attracts more migrants.

Column 4 contains our preferred specification described in Equation 7. Here we include destination-year and origin-year fixed effects to control for time and state varying variables like the price index or the presence of foreign owned firms. Despite the addition of these controls, the magnitude of the coefficient of market access decreases only slightly and remains significant at the 1% level. When excluding wages from this specification, the coefficient of market access is again higher, as it captures both the direct and indirect effect (Column 5).

Since our empirical specification derives from an aggregate discrete choice model (grouped logit model), the estimated coefficients cannot be directly interpreted as marginal effects. To find the partial effect of a change in a location characteristic on the migration probability between two states, we need to differentiate Equation 4 with respect to the X_{ij} of interest:

$$\begin{aligned}\frac{\partial s_{ij}}{\partial X_{ij}} &= \beta\frac{exp(X_{ij}\beta+\xi_{ij})[\Sigma_{j=1}^{J}exp(X_{ij}\beta+\xi_{ij})-exp(X_{ij}\beta+\xi_{ij})]}{[\Sigma_{j=1}^{J}exp(X_{ij}\beta+\xi_{ij})]^2}\\ &= \beta\frac{exp(X_{ij}\beta+\xi_{ij})}{\Sigma_{j=1}^{J}exp(X_{ij}\beta+\xi_{ij})}\frac{1-exp(X_{ij}\beta+\xi_{ij})}{\Sigma_{j=1}^{J}exp(X_{ij}\beta+\xi_{ij})}\end{aligned} \tag{14}$$

Using the empirical equivalence of the two fractions in the last part of Equation 14 (and adding again the sectoral and time dimension) simplifies this expression to

$$\frac{\partial s_{ijst}}{\partial X_{ijst}} = \beta s_{ijst}(1 - s_{ijst}) \tag{15}$$

To evaluate the importance of the direct effect of market access on domestic migration, we replace β with its estimated coefficient and s_{ijst} with the observed migration probabilities. Equation 15 then tells us how the probability of migrating from state i to any state j in sector s in year t is affected by a change in 1% in the sectoral market access gap.

The values of the elasticities for the 4183 observations in our benchmark specification (Column 4) range from 0.0003 to 0.14, depending on the observed migration probabilities. The average elasticity is 0.012, meaning that an increase by 1% in the foreign market access gap between two regions leads on average to an increase of 0.00012% in the probability of migrating for a given pair of states in sector s (in addition to any indirect effect via a higher wage gap). At first glance, this effect appears very low. However, this increase in market access by 1% leads to a growth of 34% to 57% in the number of migrants for each observation. Using the estimates from Column 5, which consider the joint effect via both channels, this increase reaches 44% to 74%.

The last two columns of Table 1 provide robustness checks. Column 6 replicates our benchmark estimation using the Poisson Pseudo-Maximum Likelihood estimator (PPML) to deal with the high number of zero-migration flows. The coefficient of our key variable of interest remains highly significant, confirming the positive impact of market access on migration rates.[28]

Column 7 addresses the concern that the positive coefficient on market access could capture state j's comparative advantage in a particular sector s. Such a comparative advantage would attract migrants in this sector regardless of its market access. To make sure that our variable of interest is indeed capturing regional differences in the access to foreign demand, we include sector-destination and sector-origin dummies.[29] A highly significant coefficient of market access also in this specification gives us further confidence in our findings.

[28] The data set in Column 6 consists of all the 1748 sector-origin-destination combinations for which we observe a positive migration flow for at least one year. The panel is not entirely balanced since we exclude 57 migration rates because i) they are constructed with less than 6 individual observations; or ii) we do not have wage data for the origin-sector combination.

[29] We exclude here the origin-year and destination-year fixed effects to reduce the number of fixed effects. Including all sets of dummies would substantially reduce the variation left to explain.

5.2 Alternative measures for market access

Before investigating any potential heterogeneous impact of market access across industries and educational levels, this section presents some sensitivity analysis on our market access measure.

Table 2: Alternative market access indicators

Dep. variable:	$\ln(\text{migrants}_{ijst}/\text{stayers}_{iist})$				
	(1) *benchmark*	(2) MA_1	(3) MA_2	(4) MA_H	(5) MA_T
$\Delta MA_{ijs\tilde{t}}$	0.571^a	0.633^a	0.560^a	0.411^a	0.254^a
	(0.097)	(0.100)	(0.102)	(0.080)	(0.031)
$\Delta w_{ijs\tilde{t}}$	0.311^a	0.320^a	0.315^a	0.316^a	0.285^a
	(0.048)	(0.048)	(0.048)	(0.048)	(0.043)
FE_{ij}	yes	yes	yes	yes	yes
FE_{st}	yes	yes	yes	yes	yes
$FE_{it}\&FE_{jt}$	yes	yes	yes	yes	yes
Observations	4183	4183	4183	4183	4183

Heteroskedasticity-robust standard errors clustered at the state of origin-year level appear in parentheses. ^{a}, ^{b} and ^{c} indicate significance at the 1%, 5% and 10% confidence levels.

To facilitate the comparison, Column 1 of Table 2 repeats our benchmark estimation from Column 4 of Table 1. The remaining columns estimate this specification using different market access indicators.

Columns 2 and 3 in Table 2 use indicators that differ in the specifications of the bilateral trade costs in the construction of market access.[30] Both indicators show positive and significant coefficients with similar magnitudes with respect to our preferred variable. Column 4 uses the foreign market access variable obtained from estimating the trade equation with the Heckman selection model (MA_H), which takes into account the presence of zero trade flows. Also here, we have a slightly lower but still positive and highly significant coefficient.

[30] See Section 3.2 for the definitions of the market access measures used in Table 2.

Finally, in Column 5, we employ an indicator for total market access, MA_T, which includes a proxy for domestic demand. Even though the aim of this study is to identify the effect of foreign demand, theoretically, access to all markets should matter for the location choice of workers and thus national and international demand should be accounted for. We expect a positive impact also for this variable, which is confirmed by our estimation.

Before concluding this section, we want to point out that our results also hold when accounting for potential sectoral reallocation of workers. As stated in Section 2, we make the implicit assumption that migration takes place only within sectors. Note that the estimation of our migration equation serves actually as a test of this hypothesis. If workers would massively switch sectors in response to changes in sectoral market access, the adjustment would take place mainly within regions. In that case, the observed migration rates could not be explained by sector-region specific variables. Consequently, the estimated coefficient of our sectoral variables would tend towards zero.

Nevertheless, in Table A-5 in Appendix A we present results on a subsample restricted to workers with sector-specific occupations. This makes it less likely that the individuals were previously working in a different sector and therefore biasing our results.[31] Even though the number of observations is reduced substantially, results hold.

The presented robustness checks confirm that our market access variable as a proxy for job opportunities linked to foreign demand is indeed an important determinant of migration patterns within Brazil. This highlights the impact of external events on the internal spatial distribution of the labor force.

[31] See Appendix B-4 for more details on the construction of this subsample. Column 1 to 4 of Table A-5 in Appendix A report results on this subsample for our benchmark specification, the PPML estimator and the indicators MA_H and MA_T.

5.3 Heterogeneous impact by industries

In this section, we discuss the sectoral dimension of our panel in more detail. Workers in different industries might react differently to changes in market access. This could arise for example from a different degree of dependence of the industries on foreign demand or different labor market structures across industries affecting the mobility of workers. To test empirically for heterogeneity in the role of market access in the migration pattern, we allow the coefficient of the market access indicator to vary across all eleven industries.

In Column 1 of Table 3, all sectors, except *Electrical* & *Electronics*, exhibit a positive and significant coefficient. This shows that the positive effect of market access that we found before is not driven by any particular sector. Column 2 also allows the coefficient on the sector-specific wage variable to vary by industry. Although this decreases the magnitudes of the market access coefficients and renders those for *Agriculture* and *Chemical* & *Pharmaceuticals* non significant, these estimates confirm the findings of Column 1.

Magnitudes of the market access coefficient vary substantially, leading to important differences in marginal effects across sectors (from on average 0.005 for *Electrical* & *Electronics* to 0.1 for *Wood*). A first indication for a possible source of such a variation across sectors lies in the sector's comparative advantage on the world market. After Brazil opened itself to foreign trade, certain sectors started to flourish, whereas others experienced a substantial decline. The industries in Table 3 are categorized into three groups (*high*, *medium* and *low*), according to their comparative advantage on the world market (for details on the classification see Appendix B-5).

Sectors with an international comparative advantage have on average higher and more significant coefficients for market access. Columns 3 and 4 repeat the estimations from

Table 3: Market access impact by sector

Dependent variable:	ln(migrants$_{ijst}$/stayers$_{iist}$)			
	(1)	(2)	(3)	(4)
High: comparative advantage industries				
$\Delta MA_{ijs\tilde{t}}$ × Agriculture	2.949^a	-0.018		
	(0.435)	(0.298)		
$\Delta MA_{ijs\tilde{t}}$ × Food	1.377^a	0.924^b		
	(0.451)	(0.424)		
$\Delta MA_{ijs\tilde{t}}$ × Wood	2.334^a	1.474^a		
	(0.433)	(0.383)		
$\Delta MA_{ijs\tilde{t}}$ × Plastic & non-metallic	0.522^a	0.234^c		
	(0.136)	(0.134)		
$\Delta MA_{ijs\tilde{t}}$ × Basic metals	1.028^a	0.529^a		
	(0.197)	(0.183)		
(β_H) $\Delta MA_{ijs\tilde{t}}$× *Strong Adv*			0.810^a	0.829^a
			(0.150)	(0.152)
Medium: no comparative advantage				
$\Delta MA_{ijs\tilde{t}}$ × Mining	1.062^b	0.915^a		
	(0.467)	(0.340)		
$\Delta MA_{ijs\tilde{t}}$ × Textiles	2.008^a	1.564^a		
	(0.241)	(0.218)		
$\Delta MA_{ijs\tilde{t}}$ × Chemical & Pharmaceuticals	0.440^b	0.263		
	(0.184)	(0.182)		
$\Delta MA_{ijs\tilde{t}}$ × Machinery and others	0.785^a	0.378^b		
	(0.153)	(0.161)		
(β_M) $\Delta MA_{ijs\tilde{t}}$ × *Medium Adv*			0.570^a	0.596^a
			(0.118)	(0.118)
Low: comparative disadvantage				
$\Delta MA_{ijs\tilde{t}}$ ×Paper & Printing	0.658^a	0.442^a		
	(0.119)	(0.113)		
$\Delta MA_{ijs\tilde{t}}$ × Electrical & Electronics	0.226	-0.383		
	(0.331)	(0.331)		
(β_L) $\Delta MA_{ijs\tilde{t}}$ × *Low Adv*			0.340^a	0.302^a
			(0.106)	(0.104)
Observations	4183	4183	4183	4183
H_0: $\beta_H = \beta_L$ (p-value)			0.001	0.000

Heteroskedasticity-robust standard errors clustered at the state of origin-year level appear in parentheses. ^{a}, ^{b} and ^{c} indicate significance at the 1%, 5% and 10% confidence levels. As in the benchmark specification of Table 2, FE_{ij}, FE_{st} and FE_{it}&FE_{jt} are included in all regressions. Columns 1 and 3 restrict the coefficient of sector-specific wage gaps to be the same across all industries. Column 2 and 4 allow the coefficient of the wage gap to vary across industries in the same way as market access. Wage coefficients are not reported for the sake of brevity. They are mostly positive and significant.

the first two columns, but restrict the coefficients so as to be the same for all industries within a group. The t-test in the bottom line of the table rejects the hypothesis of equality between the market access coefficient of the group with comparative advantage and that with a comparative disadvantage.[32]

These results suggest that workers in more international competitive industries are moving to higher market access regions and take full advantage of the positive economic prospects linked to increased exposure to exports. Our findings can thus help to explain the concentration of certain industries in specific regions.

In contrast, workers in disadvantaged industries seem less sensitive to changes in foreign market access. Since international demand for their goods is generally low, a better access to foreign markets will have less additional value for workers in these industries. As a consequence, market access is expected to play a less important role in the location decision of these workers.

6 Market access and migration by sector and education

In this final section, we distinguish between highly educated and low-educated workers. Figure A-3 in Appendix A displays differences in migrant shares between the two educational groups for each state for the years 1995 and 2003. Over the sample period, highly educated migrants were more likely to move to the South and Northeast, while the Center region has become a more popular destination for low-educated migrants. These differences in the location choices suggest that the utility of migrating to a specific state

[32] Unreported robustness checks (available upon request) show that results hold when using the various indicators for foreign market access presented in Table 2.

might vary across educational levels.

We thus investigate whether the observed differences in migration patterns can be explained partly by a heterogeneous impact of sectoral market access, depending on the educational attainment of the individuals.

However, there is no clear theoretical prediction on whether the effect of market access on migration rates should be stronger for highly educated or low-educated workers. On the one hand, a more pronounced reaction of highly qualified workers to a change in market access would be in line with the New Economic Geography model by Redding and Schott (2003). Their model predicts that higher market access leads to a higher wage premium for skilled workers. Thus, we could expect that highly educated workers have a stronger incentive to go to states with high market access to benefit from the additional wage premium or a steeper wage gradient in these regions.

On the other hand, numerous theoretical and empirical studies have suggested that highly educated workers are more sensible to certain region-specific amenities.[33] At the the same time, highly educated workers might have better access to well-paid jobs. From this perspective, higher wages and career opportunities created by a higher foreign demand could play a minor role in the migration decision of these individuals.

From the previous work by Fally et al. (2010) we know that in Brazil, the states with higher foreign market access pay low qualified workers relatively more than highly qualified workers. This finding is in line with traditional trade theory. The Stolper-Samuelson mechanism predicts that in the case of trade liberalization, there should be an increase in the relative returns of the production factor, which is relatively more

[33] For example, Levy and Wadycki (1974) have shown that in Venezuela educated individuals tend to value amenities much more than low-qualified individuals. More recently, Adamson et al. (2004) find that returns to education for the higher educated workers fall with the population size in US metropolitan areas, which is also consistent with a skill-biased effect of amenities.

abundant in the country. Gonzaga et al. (2006) confirm this prediction for Brazil, which has a relative abundance of low-skilled labor with respect to high-skilled labor. They show that wages for unskilled workers have indeed increased in Brazil during the period of trade liberalization. Thus, for low-educated workers, we could expect a strong effect of market access on migration via the indirect wage channel.

Menezes-Filho and Muendler (2011) and Corseuil et al. (2013) provide a first indication that trade liberalization could also lead to a strong adjustment via the direct channel for low-educated workers. Both studies document for Brazil that higher educational attainment contributes to increased employment durations. Low-educated workers are thus more likely to be laid off and obliged to move for a new employment.

To test for a heterogeneous role of foreign demand depending on educational attainment, we adapt Equation 7 to allow the coefficient of the independent variables to be different for highly educated and low-educated workers. Our second benchmark specification can then be written as

$$\begin{gathered}\ln m^e_{ijst} = \alpha + \beta_H \Delta MA_{ijs\widetilde{t}} \times High^e + \beta_L \Delta MA_{ijs\widetilde{t}} \times Low^e \\ + \beta_3 \Delta w^e_{ijs\widetilde{t}} \times High^e + \beta_4 \Delta w^e_{ijs\widetilde{t}} \times Low^e + FE^e_{st} + FE^e_{ij} + FE^e_{it} + FE^e_{jt} + \varepsilon^e_{ijst}\end{gathered} \tag{16}$$

where m^e_{ijst} is defined as the number of migrants in sector s belonging to educational group e in year t moving from i to j divided by the number of stayers. The dummy $High$ (Low) takes the value one when the migration rate is constructed with high(low)-educated workers. After each regression, we test for equality of β_H and β_L, to verify whether the coefficient of market access differs significantly between the two types of workers. The wage gap, $\Delta w^e_{ijs\widetilde{t}}$, is calculated using means of predicted wages that vary across states, sectors and skill groups. As before, $\widetilde{t}$ indicates that independent variables are constructed as means over the years $t-4$ to $t-1$. To take into account that other migration determinants might also vary according to educational attainment, all included fixed effects (FE^e) are allowed to differ between the two groups.[34]

6.1 Empirical results by sector and education

Table 4 reports results on the heterogeneous impact of market access across educational groups.

As in Table 1, we display first estimation results for a less restrictive specification. Column 1 does not include the state-year fixed effects, but instead the relative population size, the unemployment gap and the difference in homicide rates. Column 2 uses the same set of dummies but excludes the wage variable to obtain the joint effect of market access on migration via both channels. Column 3 contains our second benchmark specification (Equation 16) and Column 4 estimates the joint effect including all the fixed effects.

In all specifications, the coefficients of the control variables, including wages, are similar across educational groups. However, the coefficient of market access is significant at conventional levels only for low-qualified workers. The t-tests reported at the bottom of the table clearly reject the hypothesis of a uniform impact of market access across educational groups. As in the aggregate results, the magnitudes of the coefficients on market access are higher when wages are excluded from the regression, suggesting that part of the effect of a higher market access is reflected by a higher wage level.[35]

[34] This specification corresponds to splitting the sample between high and low qualified workers. Migration rates of highly educated workers represent 34% of our final sample.

[35] We confirm results of Table 4 when employing different market access measures as in Table 2 (results available upon request). Columns 5 and 6 of Table A-5 in Appendix A repeat Columns 3 and 4 of Table 4 for the subsample of workers in sector-specific occupations only. Also here we can only find a significant coefficient of market access for low-educated workers. However, only 16.5% of the observations in this subsample represent highly educated workers, therefore we do not have sufficient precision to reject the null hypothesis that impacts are the same for both types of workers.

Table 4: Bilateral migration by education

Dependent variable:	$\ln(\text{migrants}^e_{ijst}/\text{stayers}^e_{iist})$			
	(1)	(2)	(3) *benchmark II*	(4)
(β_H) $\Delta MA_{ijs\tilde{t}}$ × *High edu*	0.058	0.110	0.041	0.071
	(0.062)	(0.074)	(0.078)	(0.087)
(β_L) $\Delta MA_{ijs\tilde{t}}$ × *Low edu*	0.871[a]	0.972[a]	0.917[a]	1.069[a]
	(0.132)	(0.139)	(0.151)	(0.161)
$\Delta w^e_{ijs\tilde{t}}$ × *High edu*	0.188[a]		0.233[a]	
	(0.025)		(0.028)	
$\Delta w^e_{ijs\tilde{t}}$× *Low edu*	0.149[b]		0.202[b]	
	(0.072)		(0.090)	
$\Delta u^e_{ij\tilde{t}}$× *High edu*	-0.148[c]	-0.168[b]		
	(0.083)	(0.082)		
$\Delta u^e_{ij\tilde{t}}$× *Low edu*	-0.167[c]	-0.172[c]		
	(0.095)	(0.094)		
$\Delta pop_{ij\tilde{t}}$× *High edu*	1.073	1.470		
	(0.937)	(0.956)		
$\Delta pop_{ij\tilde{t}}$× *Low edu*	-0.184	-0.153		
	(0.996)	(0.961)		
$\Delta death_{ij\tilde{t}}$× *High edu*	-0.169	-0.158		
	(0.118)	(0.114)		
$\Delta death_{ij\tilde{t}}$× *Low edu*	-0.043	-0.041		
	(0.069)	(0.068)		
FE^e_{ij}	yes	yes	yes	yes
FE^e_{st}	yes	yes	yes	yes
$FE^e_{it}\&FE^e_{jt}$	no	no	yes	yes
Observations	4614	4614	4614	4614
H_0: $\beta_H = \beta_L$ (p-value)	0.000	0.000	0.000	0.000

Heteroskedasticity-robust standard errors clustered at the state of origin-year level appear in parentheses. [a], [b] and [c] indicate significance at the 1%, 5% and 10% confidence levels.

In light of our findings, the differences in the observed migration patterns across educational levels can be partly explained by a different sensitivity to foreign market access.

Economic opportunities associated with international trade seem most important for the location choice of low educated individuals. The strong impact of market access for low-qualified workers can be explained by the fact that Brazil is exporting mainly goods that are intensive in unskilled labor. Consequently, an increase in the demand for exported goods signifies a higher demand and more jobs for low-educated workers. Also when controlling for wage differentials, β_L remains highly significant. This indicates that new employment opportunities created by a stronger local export activity are indeed important for the location choice of this group.

For the highly educated workers, the market access indicator remains insignificant in all specifications, even if we exclude wages in order to estimate the joint impact of foreign market access via both channels. The interpretation of this result is less straight forward since it might be driven by various forces, as explained above.

One possible explanation is that these individuals have in general easier access to "high quality" jobs with good working conditions and career prospects. The alternative explanation of a predominant role of amenities with respect to economic considerations is however at odds with the fact that highly educated workers are also responsive to wage differentials.

Before we conclude, we use the estimated coefficients of Column 3 of Table 4 to simulate the implied change in each observed migration rate in response to a positive shock in the foreign demand for Brazilian goods. This provides more intuition for our results and allows to identify the regions that are particularly affected by a specific demand shock.

In Table A-6 in Appendix A, we simulate the effects of four different shocks to $\Delta MA_{ijs\tilde{t}}$. Column 1 reports the average share of immigrants of each state over the sample period. Columns 2 to 5 show for each state how this share would be affected by these different demand shocks. In Columns 2 to 4, we study an increase of the demand for Brazilian goods from specific groups of countries. In Column 5, we decrease bilateral trade costs.

Marginal effects of market access for the highly educated workers being very low, the implied change in the number of migrants is driven by the low-educated workers. Note that the numbers presented in this table correspond to partial equilibrium effects since our simulation rules out any impact of market access on migration going through the indirect effect of wage differentials. Also, the model does not incorporate any potential impacts of migration on housing costs or other congestion costs.[36]

The first scenario supposes an increase of 3% in the demand coming from the Mercosur members (Argentina, Uruguay and Paraguay).[37] This increase in the relative importance of the Mercosur countries affects states and sectors differently. Notably the increase in market access of the Southern states which are closer to the Mercosur partners will be higher than for the Northern states. The resulting change in the market access gap between two states impacts directly the bilateral migration rates. When summing over all sending states, sectors and the two educational groups, we can calculate the total number of additional immigrants a state will receive. Column 2 shows that states in the South will see the most important increase in their share of immigrants, whereas the North and Northeast are relatively less well connected to these markets and will attract less migrants.

[36] However, the additional effects mentioned here should play only a minor role. Notably, Morten and Oliveira (2014) show that congestion costs associated to housing would be negligible in the case of Brazil.

[37] This positive demand shock is modeled as an increase by 3% of the market capacity (the estimated importer fixed effect FM_{jst} in Equation 12) of these countries. An increase of 3% corresponds to an increase by one standard deviation of the estimated market capacities.

Columns 3 and 4 repeat the exercise for an increase in 3% of the demand coming from one of the other two main destinations of Brazil's exports, respectively the EU and the NAFTA countries. Results are different here: for these two scenarios it is the Southern states that will see the strongest decrease in the share of immigrants. In contrast, the geographic proximity of the Northern and the Northeastern states to the EU and NAFTA countries leads to an important increase in their market access and hence in the immigration share of these states.

These changes in migration patterns in response to a change in the access to foreign demand illustrates well, how much the spatial structure of the domestic economy is influenced by what happens abroad.

In the last scenario (Column 5), we consider a decrease in the internal distance to the next harbor by 10%, i.e a reduction in bilateral trade costs (ϕ_{ij}). This improves relatively more foreign market access of inland states.

This last finding has also implications for domestic policies: the reduction in internal distance can also be interpreted as an improved domestic infrastructure that facilitates the access to the sea for the inland states. However, as our results point out, policy makers aiming at regional development need to be aware that due to the country's integration into the world economy the effects of their measures can be reinforced or opposed by events happening outside of the country.

7 Conclusion

The main objective of this paper is to gain a better understanding of how international trade affects the spatial distribution of economic activity inside the country. We do so by examining the effect of regional and sectoral variations in the access to foreign demand on bilateral migration rates between Brazilian states.

An increase in a region's access to foreign demand represents a higher local labor demand. We argue that a higher market access attracts migrants via two channels: i) an indirect effect via an increase in the local wage level, and ii) a direct impact via new job opportunities. This direct effect can be driven by a higher number of vacancies or improved job characteristics such as better working conditions or career prospects that arise after an increased foreign demand.

This paper shows that workers indeed move away from states with low market access and prefer states with higher market access. By controlling for region and sector-specific wages, we can identify the direct impact that market access has on the migration decision beyond the wage channel.

We further find differences in the sensitivity of migration rates to changes in foreign demand across sectors and educational levels. This heterogeneity can reinforce the industrial specialization of regions and explain differences in migration patterns between groups of workers.

Our findings highlight the importance of interactions with foreign countries in shaping the internal spatial distribution of the labor force. This aspect is generally excluded from regional migration studies, which rely only on purely domestic migration determinants. In contrast to classical migration determinants such as spot wages or unemployment rates, our market access measure also captures long term considerations, such as career prospects.

This paper employs household survey data, which has the advantage of considering the informal sector that represents over one third of the Brazilian workforce. However, our data doesn't allow us to identify the main driving force behind the observed direct effect of foreign market access. In order to better understand the role of international trade on internal migration, more work is needed to identify the relative importance of additional job openings and regional differences in job characteristics. Linked employer-employee data could be used to study the evolution of the wage profile after migration or to assess non-pecuniary aspects of the jobs (e.g. job tenure) and how they are linked to the export activity of a region.

References

Adamson, D., D. Clark, and M. Partridge (2004). Do urban agglomeration effects and household amenities have a skill bias? *Journal of Regional Science 44*(2), 201–223.

Aguayo-Tellez, E., M.-A. Muendler, and J. P. Poole (2010). Globalization and formal sector migration in Brazil. *World Development 38*(6), 840–856.

Anderson, J. E. and E. van Wincoop (2003). Gravity with gravitas: A solution to the border puzzle. *American Economic Review 93*(1), 170–192.

Autor, D. H., D. Dorn, and G. H. Hanson (2013). The China syndrome: Local labor market effects of import competition in the United States. *American Economic Review 103*(6), 2121–2168.

Berry, S. T. (1994). Estimating discrete-choice models of product differentiation. *RAND Journal of Economics 25*(2), 242–262.

Bertoli, S. and J. Fernández-Huertas Moraga (2013). Multilateral resistance to migration. *Journal of Development Economics 102*(C), 79–100.

Bertoli, S., J. Fernández-Huertas Moraga, and F. Ortega (2013). Crossing the border: Self-selection, earnings and individual migration decisions. *Journal of Development Economics 101*(C), 75–91.

Chaney, T. (2008). Distorted gravity: The intensive and extensive margins of international trade. *American Economic Review 98*(4), 1707–21.

Corseuil, C. H., M. Foguel, G. Gonzaga, and E. P. Ribeiro (2013). Youth labor market in Brazil through the lens of the flow approach. Proceedings of the 41st brazilian economics meeting, ANPEC.

Crozet, M. (2004). Do migrants follow market potentials? An estimation of a New Economic Geography model. *Journal of Economic Geography 4*(4), 439–458.

Cunha, J.-M. P. d. (2002). O uso das PNADs na análise do fenômeno migratório: Possibilidades, lacunase desafios metodológicos [Using PNAD surveys to study migrations: Potentialities, shortcomings and methodological challenges]. IPEA working paper 875.

Dahl, G. (2002). Mobility and the return to education: Testing a Roy model with multiple markets. *Econometrica 70*(6), 2367–2420.

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

Fally, T., R. Paillacar, and C. Terra (2010). Economic geography and wages in Brazil: Evidence from micro-data. *Journal of Development Economics 91*(1), 155–168.

Federação das Indústrias do Estado do Rio de Janeiro FIRJAN (2010). Quanto custa abrir uma empresa no Brasil? Estudos para o desenvolvimento do Estado do Rio de Janeiro, no. 6, Rio de Janeiro.

Fiess, N. and D. Verner (2003). Migration and human capital in Brazil during the 1990s. World Bank Policy Research Working Paper 3093, World Bank.

Fujita, M., P. Krugman, and A. Venables (1999). *The Spatial Economy: Cities, Regions, and International Trade*. Cambridge: MIT Press.

Goldberg, P. and N. Pavcnik (2007). Distributional effects of globalization in developing countries. *Journal of Economic Literature 45*(1), 39–82.

Gonzaga, G., N. Menezes-Naércio, and C. Terra (2006). Trade liberalization and the evolution of skill earnings differentials in Brazil. *Journal of International Economics 68*(2), 345–367.

Grogger, J. and G. H. Hanson (2011). Income maximization and the selection and sorting of international migrants. *Journal of Development Economics 95*(1), 42–57.

Hanson, G. (2005). Market Potential, Increasing Returns, and Geographic Concentration. *Journal of International Economics 67*(1), 1–24.

Hausman, J. and D. McFadden (1984). Specification tests for the multinomial logit model. *Econometrica 52*(5), 1219–1240.

Head, K. and T. Mayer (2006). Regional wage and employment responses to market potential in the EU. *Regional Science and Urban Economics 36*(5), 573–594.

Head, K. and T. Mayer (2011). Gravity, market potential and economic development. *Journal of Economic Geography 11*(2), 281–294.

Head, K. and T. Mayer (2013). Gravity equations: Workhorse, toolkit, and cookbook. CEPR Discussion Papers 9322, C.E.P.R. Discussion Papers.

Helpman, E., O. Itskhoki, M.-A. Muendler, and S. Redding (2012). Trade and inequality: From theory to estimation. CEP Discussion Papers dp1138, Centre for Economic Performance, LSE.

Helpman, E., M. Melitz, and Y. Rubinstein (2008). Estimating Trade Flows: Trading Partners and Trading Volumes. *Quarterly Journal of Economics 123*(2), 441–487.

Henley, A., G. R. Arabsheibani, and F. G. Carneiro (2009). On defining and measuring the informal sector: Evidence from Brazil. *World Development 37*(5), 992–1003.

Hering, L. and S. Poncet (2010). Market access impact on individual wages: Evidence from China. *Review of Economics and Statistics. 92*(1), 145–159.

Kovak, B. (2011). Local labor market effects of trade policy: Evidence from Brazilian liberalization. mimeo.

Kovak, B. (2013). Regional efects of trade reform: What is the correct measure of liberalization? *American Economic Review 103*(5), 1960–1976.

Krishna, P., J. P. Poole, and M. Z. Senses (2014). Wage effects of trade reform with endogenous worker mobility. *Journal of International Economics*, forthcoming.

Letouzé, E., M. Purser, F. Rodríguez, and M. Cummins (2009). Revisiting the migration-development nexus: A gravity model approach. MPRA Paper 19227, University Library of Munich, Germany.

Levy, M. B. and W. J. Wadycki (1974). Education and the decision to migrate: An econometric analysis of migration in Venezuela. *Econometrica 42*(2), 377–88.

McFadden, D. (1974). *Frontiers of Econometrics*, Chapter Conditional Logit Analysis of Qualitative Choice Behavior, pp. 105–142. Academic Press.

Menezes-Filho, N. A. and M.-A. Muendler (2011). Labor reallocation in response to trade reform. NBER Working Papers 17372, National Bureau of Economic Research, Inc.

Ministry of External Relations, Brazil (2008). Brazilian ports: a safe haven for international investment. Technical report, Trade Promotion department.

Morten, M. and J. Oliveira (2014). Migration, roads and labor market integration: Evidence from a planned capital city. mimeo.

Muendler, M.-A. (2007). Balassa (1965) comparative advantage by sector of industry, brazil 1986-2001. mimeo.

Muendler, M.-A. (2008). *The Analysis of Firms and Employees: Quantitative and Qualitative Approaches*, Chapter Trade and Workforce Changeover in Brazil. The University of Chicago Press.

Muriel, B. and C. Terra (2009). Sources of comparative advantage in Brazil. *Review of Development Economics 13*(1), 15–27.

Ortega, F. and G. Peri (2013). The effect of income and immigration policies on international migration. *Migration Studies 1*(1), 47–74.

Redding, S. and P. K. Schott (2003). Distance, skill deepening and development: Will peripheral countries ever get rich? *Journal of Development Economics 72*(2), 515–541.

Redding, S. and A. Venables (2004). Economic geography and international inequality. *Journal of International Economics 62*(1), 53–82.

Topalova, P. (2010). Factor immobility and regional impacts of trade liberalization: Evidence on poverty from India. *American Economic Journal: Applied Economics 2*(4), 1–41.

Volpe-Martincus, C. (2010). Spatial effects of trade policy: Evidence from Brazil. *Journal of Regional Science 50*(2), 541–869.

Wagner, J. (2012). International trade and firm performance: a survey of empirical studies since 2006. *Review of World Economics 148*(2), 235–267.

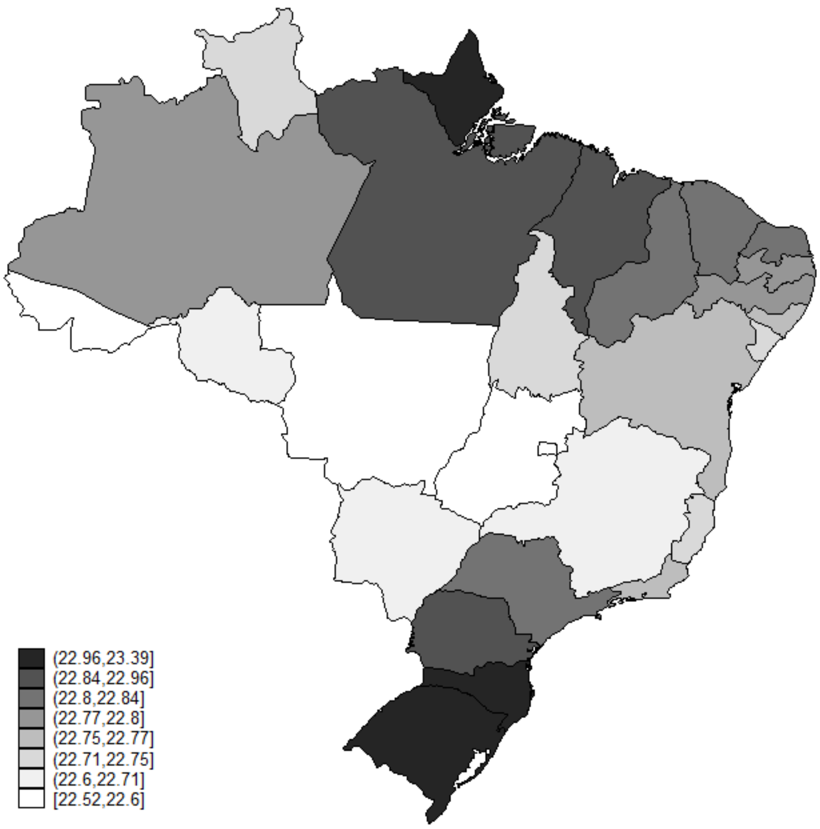
World Bank (2009). *World Development Report 2009: Reshaping Economic Geography*. World Bank.

Appendix A - Additional tables and figures

Figure A-1: Map of Brazil

Source: http://volunteerbrazil.com/photos%202008/map-of-brazil.gif

Figure A-2: Average market access by state (over time and sectors)



Source: own calculations. The figure shows the spatial distribution of the average market access over our sample period. States are grouped into eight quantiles.

Table A-1: Estimation results of the trade equation (Equation 11): Averages of coefficients by industry

Industry	Distance	Border	RTA	WTO	Market capacity
Agriculture	-1.134	1.058	0.549	0.417	26.44
	(0.0455)	(0.245)	(0.127)	(0.248)	(1.777)
Mining	-1.457	1.040	0.317	0.230	27.19
	(0.0393)	(0.0885)	(0.0972)	(0.146)	(1.944)
Food	-0.739	0.726	0.289	0.183	18.66
	(0.0899)	(0.171)	(0.199)	(0.373)	(2.017)
Textiles	-1.334	1.078	0.392	0.417	28.82
	(0.0155)	(0.229)	(0.119)	(0.211)	(1.760)
Wood	-1.401	0.584	0.778	0.312	28.97
	(0.0315)	(0.187)	(0.167)	(0.178)	(2.021)
Paper & Printing	-1.689	0.735	0.727	0.503	30.74
	(0.0631)	(0.151)	(0.0801)	(0.205)	(1.866)
Chemical & Pharmaceuticals	-1.520	0.577	0.608	0.346	30.78
	(0.0191)	(0.149)	(0.0944)	(0.119)	(1.861)
Plastic & non-metallic	-1.616	0.768	0.619	0.560	30.75
	(0.0309)	(0.144)	(0.171)	(0.180)	(1.687)
Basic metals	-1.572	0.589	0.558	0.375	31.04
	(0.0330)	(0.187)	(0.142)	(0.166)	(2.028)
Electrical & Electronics	-1.398	0.596	0.580	0.734	30.14
	(0.0287)	(0.200)	(0.186)	(0.212)	(1.914)
Machinery	-1.430	0.799	0.546	0.660	31.25
	(0.0319)	(0.124)	(0.157)	(0.201)	(1.740)

Equation 11 is run separately for every industry-year combination. This corresponds to 12 regressions for each industry. This table shows averages of coefficients by industry. Standard deviations of the coefficients are indicated in parentheses.

Table A-2: Summary Statistics

	Obs.	Mean	Std. dev.	Min.	Max.
m_{ijst}	4183	0.0277	0.054	0.0004	0.824
$ln(m_{ijst})$	4183	-4.587	1.401	-7.873	-0.194
$\Delta pop_{ij\tilde{t}}$	4183	-0.186	1.394	-4.802	4.210
$\Delta death_{ij\tilde{t}}$	4183	0.0284	0.949	-2.407	2.407
$\Delta u_{ij\tilde{t}}$	4183	-0.0435	0.574	-1.919	1.668
$\Delta w_{ijs\tilde{t}}$	4183	-0.0050	0.809	-4.656	3.703
$\Delta MA_{ijs\tilde{t}}$	4183	-0.0075	0.232	-1.572	1.540
ΔMA_1	4183	0.0177	0.193	-1.299	1.398
ΔMA_2	4183	-0.0108	0.230	-1.525	1.470
ΔMA_T	4183	-0.0302	0.916	-3.169	3.203
ΔMA_H	4183	-0.0121	0.280	-2.025	1.965

Table A-3: Correlation table

	$ln(m_{ijst})$	$w_{ijs\tilde{t}}$	$MA_{ijs\tilde{t}}$	MA_1	MA_2	MA_T	MA_H
$ln(m_{ijst})$	1.000						
$\Delta w_{ijs\tilde{t}}$	0.225^a	1.000					
$\Delta MA_{ijs\tilde{t}}$	0.079^a	-0.003	1.000				
ΔMA_1	0.037^b	-0.058^a	0.791^a	1.000			
ΔMA_2	0.039^b	-0.075^a	0.988^a	0.753^a	1.000		
ΔMA_T	0.334^a	0.555^a	0.231^a	0.052^a	0.141^a	1.000	
ΔMA_H	0.055^a	-0.037^b	0.992^a	0.783^a	0.989^a	0.203^a	1.000
Observations	4183						

^{a}, ^{b} and ^{c} indicate significance at the 1%, 5% and 10% confidence levels.

Table A-4: Migration rates by sectors

Industry	Migration rates	Nb of individuals
	Averages over all years	
Agriculture	2.66	16026.50
Mining	3.37	414.00
Food	3.54	3402.88
Textiles	2.80	5376.13
Wood	4.14	1043.00
Paper & Printing	2.78	966.13
Chemical & Pharmaceuticals	3.54	1075.25
Plastic & non-metallic	3.13	1621.63
Basic metals	2.87	3414.13
Electrical & Electronics	3.04	583.50
Machinery	2.89	1567.50

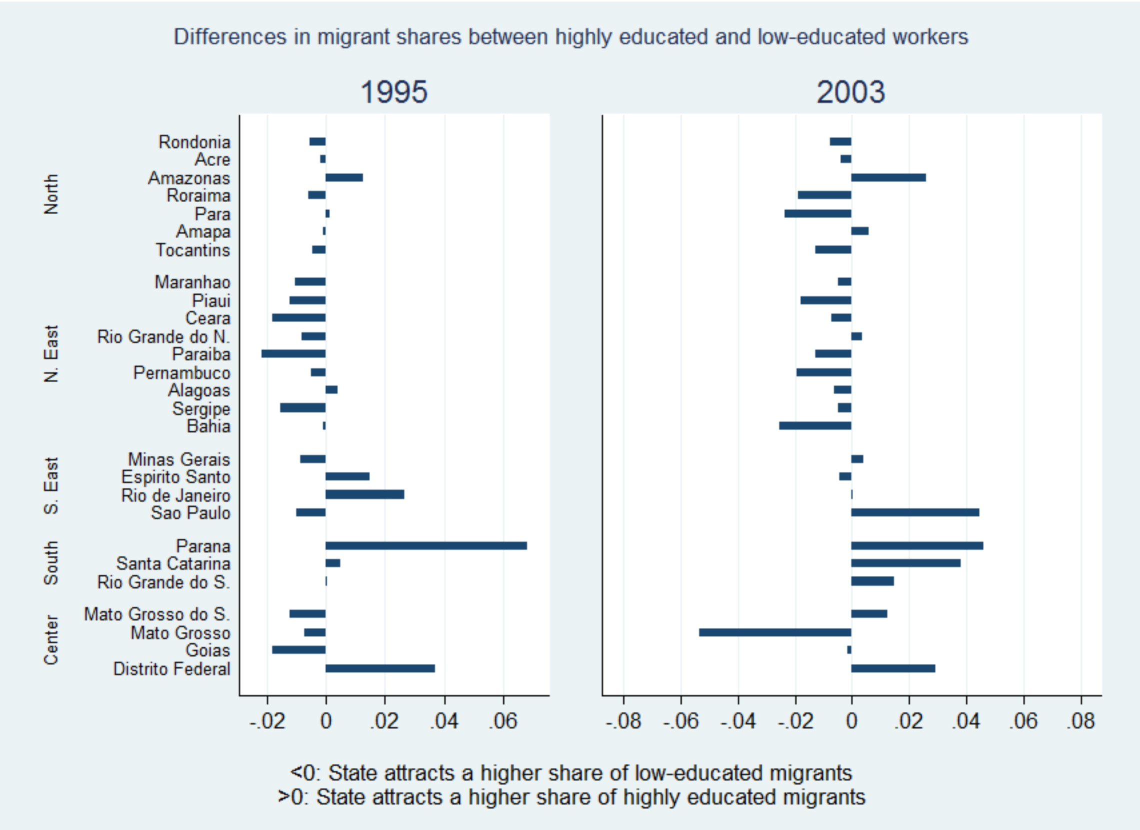
Source: Own calculations. Data are from the PNAD (1995 - 2003).

Table A-5: Main results for sector-specific occupations only

Dependent variable:	$\ln(\text{migrants}_{ijst}/\text{stayers}_{iist})$				$\ln(\text{migrants}^e_{ijst}/\text{stayers}^e_{iist})$	
	(1) *benchmark I*	(2) PPML	(3) MA_H	(4) MA_T	(5) *benchmark II*	(6)
$\Delta MA_{ijs\tilde{t}}$	0.443^b	0.636^c	0.327^b	0.259^a		
	(0.189)	(0.374)	(0.154)	(0.038)		
$\times$*High edu*					-2.608	-1.254
					(2.128)	(1.777)
$\times$*Low edu*					0.519^a	0.790^a
					(0.195)	(0.223)
$\Delta\ w^e_{ijs\tilde{t}}$	0.361^a	0.184^c	0.367^a	0.322^a		
	(0.105)	(0.100)	(0.104)	(0.098)		
$\times$*High edu*					0.365	
					(0.313)	
$\times$*Low edu*					0.428^a	
					(0.132)	
Observations	2181	6500	2181	2181	2392	2392

Heteroskedasticity-robust standard errors clustered at the state of origin-year level appear in parentheses. ^{a}, ^{b} and ^{c} indicate significance at the 1%, 5% and 10% confidence levels. Columns 1 to 4 include the fixed effects $FE_{it}, FE_{jt}, FE_{st}$ and FE_{ij}. Columns 5 and 6 include fixed effects for $FE^e_{it}, FE^e_{jt}, FE^e_{st}$ and FE^e_{ij}.

Figure A-3: Differences in migration patterns across educational groups



Source: Own calculations. Migrant shares for each state are calculated as migrants from i to j with educational level e over the total number of migrants in the respective educational group.

Table A-6: Effects of changes in MA

	State	Immigrants	positive demand shock in			Decrease in
		(share in local pop.) (in %) (1)	Mercosur (in %) (2)	EU (in %) (3)	NAFTA (in %) (4)	internal distance (in %) (5)
North	Rondonia	.8	-6.5	1.7	5	-.2
	Acre	1.4	-.8	.8	.3	1.3
	Amazonas	.7	-1.7	-.6	.6	.9
	Roraima	4.2	-1.1	-1.1	1.5	1.2
	Para	1	-1.3	-.6	2.2	-.3
	Amapa	.4	-.6	-1.4	1.8	-.2
	Tocantins	6.7	-1.8	.9	1.2	1
Northeast	Maranhao	.3	-1.6	.8	1.3	-.2
	Piaui	1.3	-3	1.7	2	.1
	Ceara	1	-3.4	2.9	1.1	-.1
	Rio Grande do N.	1.2	-4.4	2.9	1.8	-.6
	Paraiba	.8	-4.1	3.5	.6	0
	Pernambuco	.6	-3.6	3.1	.4	0
	Alagoas	.4	-1	.8	.1	-.1
	Sergipe	1.5	-3.2	1.9	1.1	.1
	Bahia	.5	-2.7	1.8	.6	-.3
Southeast	Minas Gerais	.8	-.7	.3	0	-.2
	Espirito Santo	1.2	-.7	.5	-.2	-.2
	Rio de Janeiro	.2	1.3	-.9	-.6	-.4
	Sao Paulo	.6	3.3	-1.9	-1.4	-1
South	Parana	1.2	5.6	-2.1	-2	6.4
	Santa Catarina	1.1	2.8	-.9	-1.2	-3.9
	Rio Grande do S.	.7	11.4	-5.1	-5.1	-2.2
Center	Mato Grosso do S.	3.8	-1.9	.4	.4	-.3
	Mato Grosso	3.9	-1.2	.1	-.5	.3
	Goias	1.6	.8	-.5	-1.2	.2
	Distrito Federal	4	1.4	-.8	-1.5	.6

Source: own calculations. Immigrant shares in Column 1 are the observed shares in the PNAD, constructed based on the sectors included in our analysis. The changes in the immigration shares in Columns 2 to 5 are obtained with help of the estimates of Column 3 of Table 4. Column 2 to 4 simulate the consequences of an increase by 3% in the market capacity of the corresponding group of countries. Column 5 assumes a decrease in the internal distance by 10%.

Appendix B - Data sources and classifications

B-1 Industry classification

The industrial classification used in the PNAD is PNAD/CD91 for the years 1992 to 2001 and CNAE 1.0 for 2002 and 2003. Using the correspondence tables proposed by the Brazilian National Commission CONCLA (IBGE), we can map these industry codes into the ISIC Rev 3 classification at the two digit level. However, to assure that all categories are well-matched over the entire period, we further need to aggregate up to 12 sectors. We discard *Petroleum and Gas* because of a very low number of workers in this sector. Our final sample consists of 11 tradeable industries (see Table A-1 for the complete list).

B-2 Trade flows and trade costs variables

In order to estimate sector-specific market access we need two sets of trade data for the years 1991 to 2002: (1) trade data between Brazilian states and foreign countries (Secretaria de Comercio Exterior, Ministry of Trade); and (2) between foreign economies (BACI: Base pour l'analyse du Commerce International, CEPII). The original industry classification in these datasets corresponds to the HS-2 level (Brazilian exports and imports) and the HS-6 level (BACI). These are matched into the eleven industrial sectors defined in the migration data.

The data for RTA, WTO, *common border* and *landlocked* are provided by the CEPII. For the construction of MA_H we estimate the trade equation using a two-stage Heckman estimation procedure. As selection variable, we use a *Doing Business* indicator, which estimates the costs of starting a business (measured as % of GDP). For the foreign countries, we use the information on this variable from the World Bank. For the Brazilian states, the data comes from the Federação das Indústrias do Estado do Rio de Janeiro FIRJAN (2010). By multiplying the value of this indicator of the two trading partners, we obtain a bilateral measure for every exporter-importer combination.

B-2.1 Bilateral Distances

Distances between two foreign countries are defined as the geodesic distance between the two largest cities (provided by the CEPII). Bilateral distance between Brazilian states and foreign countries are the sum of an internal component and an external component.

The internal component is the average geodesic distance of the state's cities with more than 100,000 inhabitants to their nearest harbor.[38] The external component is the mean of the geodesic distances between these nearest harbors and the main city of the partner country.

We choose to incorporate the internal distance to the nearest harbor since about 95% of the volume of Brazilian exports leave the country via the sea (Ministry of External Relations, Brazil, 2008). Transport modes via land, air and rail being much more expensive, a location far away from a harbor increases trade costs substantially, which is captured by the higher internal distance. The market access measure MA_1 uses the great circle distance between the Brazilian states and the foreign country leads. This leads to highly similar results for both the estimates in the trade equation and the migration equation (Column 2 of Table 2).

The construction of the total market access variable (MA_T) uses also the bilateral distances between the Brazilian states. These are calculated as geodesic distances between their respective capitals.

The harbors used to calculate the shortest internal distance are: Maceio (AL)[39], Macapa (AP), Ilheus (BA), Aratu (BA), Pecem (CE), Fortaleza (CE), Vitoria (ES), Itaqui (MA), Antonina (PA), Belem (PA), Cabedelo (PB), Suape (PE), Paranagua (PR), Angra dos Reis (RJ), Niteroi (RJ), Natal (RN), Porto Alegre (RS), Rio Grande (RS), Imbituba (SC), Itajai (SC), San Sebastiao (SP) and Santos (SP). The coordinates of the ports are taken from http://www.worldportsource.com/ports/BRA.php.

B-3 Control variables in the migration equation

Data on homicide rates and population of the Brazilian states come from the IPEA (*Instituto de Pesquisa Econômica Aplicada*). Since official unemployment data at the state level is not available, we compute yearly unemployment rates at the state level with the information of the employment status of the total working age population in the PNAD. In Section 6, we use unemployment rates that are constructed separately for each educational group.

[38] For states with numerous cities with more than 100,000 inhabitants, we take the largest five cities. Data for city population in 2000 and longitudes and latitudes comes from http://www.world-gazetteer.com/.

[39] Abbrevations in parentheses correspond to the official acronyms of the Brazilian states.

B-4 Occupations

Robustness checks in Table A-5 use a subsample of workers with industry-specific occupations. The occupational classifications used in the PNAD changes in 2002. This robustness check is thus limited to the years 1995 to 2001, for which the classification remains the same.

The PNAD contains ten wide occupational groupings (e.g. administrative, scientific & high tech, retail, manufacturing). These can be broken down into in total 381 occupations. Our subsample includes only workers in occupations which can be mapped easily into our ISIC industries. We also use the concordance tables between the national classification in the PNAD and the International Standard Classication of Occupations (ISCO-88) to confirm our match and to drop further occupations, which are not sufficiently industry-specific. This leaves us with 156 occupations for which we can safely argue that they are industry-specific enough to assume that workers remain in the same industrial sector also when changing employment (e.g. *goldsmiths* or *clock and watch repairs/makers*).

B-5 Comparative advantage of industries

To determine whether an industry enjoys a comparative advantage or disadvantage on the international market, we employ the measure of comparative advantage for Brazilian industries proposed by Muendler (2007). He computed the Balassa indicator of revealed comparative advantage for every industry at the ISIC 2 level. The Balassa indicator is defined as the ratio of the export share of a specific industry in Brazil over the average export share of the same industry at world level. A sector with a ratio higher than 1 is said to have a Revealed Comparative Advantage (RCA). As this differentiation may be considered too strict, it is safer to use the distribution of the RCA indicator. Muendler reports the distribution in quintiles and comments that industries are stable across quintiles during the period 1990-97.

We classify the different industries according to their comparative advantage into three groups: $High$, $Medium$ and Low. The Balassa indicator industries in the comparative advantage group ($High$) lie in the fifth or fourth quintile of the distribution, while the comparative disadvantage group (Low) lie in the lowest two quintiles. The remaining industries are considered as not having any particular comparative advantage ($Medium$).