The (Rail)road to Structural Change: Transportation Costs, Integration and Production Specialization in a Regional Economy

Sarah Walker ∗
University of Wisconsin, Madison

December 29, 2012
Preliminary Draft: DO NOT CITE

Abstract

The current paper examines whether or not reductions in transportation costs characterize long-run patterns of production specialization in a regional economy, and the possible welfare effects of these patterns on price convergence and wage dispersion. In answering this question I exploit a natural experiment in the Austro-Hungarian (or "Habsburg") Empire, where massive railroad development occurred throughout the latter half of the 19th century and into the early 20th century. Summary statistics suggest that throughout this period regional production patterns emerged, with some regions specializing in industrial production and others in more agrarian goods. To provide intuition, I sketch a 2 x 2 Ricardian model of trade with transportation costs and examine trends in regional specialization patterns, prices, and wages as a result of decreasing transportation costs. The paper contributes to the literature on infrastructure and economic development by using an instrumental variables approach to addressing simultaneity biases inherent in these estimations. The results show strong support for price convergence and wage dispersion as a result of decreasing transportation costs, and moderate support for regional specialization in industrial goods.

KEYWORDS: transportation costs, railroads, infrastructure, specialization, Ricardian trade, market integration, Eastern Europe, nineteenth century economic development.

∗3rd Year PhD student in the Agricultural and Applied Economics Department at the University of Wisconsin, Madison. Email: slwalker5@wisc.edu
1 Introduction

From Heckscher-Olin to the 3-factor, n-good cones of diversification, structural change is considered an inevitable part of economic development. Similarly, the growing stock of literature seeking to measure the effects of economic integration on growth and welfare — while ambiguous in proving overwhelming gains — often concludes that a relative degree of openness is essential to economic development (Estevadeordal and Taylor, 2008; Bagwhati, 2004; Stiglitz, 2002). Yet, if structural change is essential to growth and trade enhances it, what, then, is the relationship between structural change and trade?

In the structural change literature, it seems to be a foregone conclusion that production factors move. Meanwhile, there is a dearth of research dedicated to quantifying the movement of these factors. To the extent that trade allows goods and services to flow between economies, it seems intuitive that reductions in trade barriers would be a natural catalyst to structural change and, hence, growth. Ventura (1997) shows that international factor-mobility is crucial to overcoming decreasing returns-to-scale, while McMillan and Rodrik (2011) empirically illustrate that the direction of factor movements is essential to sustained economic growth. But what is the catalyst of these factor movements?

The current study seeks to explore the role of economic integration in catalyzing structural change. Specifically, it examines whether or not reductions in transportation costs characterize long-run patterns of production specialization within a regional economy, and the possible welfare effects of these patterns on price convergence and wage dispersion. In answering this question I exploit a natural experiment in the Austro-Hungarian (or ”Habsburg”) Empire, where massive railroad development occurred throughout the latter half of the 19th century and into the early 20th century. Summary statistics suggest that throughout this period regional production patterns
emerged, with some regions specializing in industrial production and others in more agrarian goods. To provide intuition, I sketch a 2 x 2 Ricardian model of trade with transportation costs and examine trends in regional specialization patterns, prices, and wages as a result of decreasing transportation costs. As empirical support, I run a series of two-step instrumental variables regressions that estimate the effect of increased railroad construction on price convergence, wage dispersion, and regional specialization in production.

Section 2 provides a brief review of the literature on structural change and trade, while section 3 discusses the Habsburg natural experiment. Sections 4 and 5 present the theoretical model and empirical strategy, respectively, while Section 6 describes the data. Section 7 provides the preliminary estimates and discussion, and Section 8 concludes.

2 Previous Discussions of Infrastructure, Integration, and Structural Change

2.1 Infrastructure, Integration and the Impact on Economic Outcomes

There is a growing body of work dedicated to understanding the impact of massive infrastructure developments on economic integration, development, and production patterns (Keller and Shiue, 2008; Donaldson, 2010; Banerjee et al., 2012; Rothenberg, 2012; Storeygard, 2012). Keller and Shiue (2008) explore the effect of tariff reductions versus decreased transportation costs in stimulating trade in the 19th century German empire—a context that is geographically and historically quite relevant to the current study—and find that the introduction of railroads had a more significant effect on decreasing price dispersion than reductions in tariffs (Keller and Shiue, 2008). Donaldson’s (2010) work on railroads in 19th century India is consistent with these finding and suggests an even wider range of effects of decreased transportation costs through decreased price dispersion, promotion of interregional trade and increased incomes (Donaldson, 2010).

Nonetheless, there is an existing debate over the role that infrastructure plays in
economic growth. In a similar fashion to Fogel’s (1962) work concerning the effect of the American railroad on economic growth, Banerjee et al. (2012) find that proximity to transportation networks in China have a moderate effect on per capita GDP levels and no effect on per capita GDP growth. In addition, access to transportation infrastructure may increase inequality (Banerjee et al., 2012). The patterns of inequality are particularly interesting when examined through the relationship between infrastructure improvements and production location decisions. The New Economic Geography (NEG) literature (Krugman 1991; Krugman and Venables, 1995; Fugita and Venables, 1999; Puga, 1999) provides formal models to explain the mechanisms through which reductions in transportation costs may drive regional inequalities (i.e., ”core” and ”periphery” patterns), namely via wage differentials that affect the location decisions of firms across space. Recently, a series of empirical applications of the traditional NEG framework is now beginning to shed quantitative light on these predictions. Rothenberg (2012) examines road improvements in Indonesia and finds that the resulting reductions in transportation costs are accompanied by a dispersion of manufacturing firms. While he does not test the welfare implications of this phenomenon, there are clear potential effects, as road improvements affected rural and urban regions differently. Other direct tests of the NEG theory also find support for the relationship between transportation costs and regional inequality (Brakman, 2005; Redding and Venables, 2004); although the emphasis of these studies is typically on modern Western Europe. In addition, many of these empirical studies focus on relatively short periods of time, Therefore, the present study contributes to this growing literature by examining long-run patterns (over 70 years) of massive infrastructure developments and their consequent effects on production patterns, regional specialization, and long-run economic outcomes.
2.2 Economic Integration and Structural Change

The underlying argument of the current study is that economic integration is the necessary mechanism through which long-term structural change occurs (via its ability to incentivize the movement of factors). Hence, eliminating barriers to trade (i.e., tariffs, transportation costs and restrictions on the mobility of factors) should catalyze observable patterns in production. To the extent that structural change induces growth, one should find that structural change-inducing reductions in trade barriers should also result in welfare gains. While there are few studies that link trade barriers to structural change directly, a rich body of work exists that examines the effects of trade barrier reductions on economic growth and welfare.

Estevadeordal and Taylor (2008) confront the heated debate surrounding the effects of trade policy on growth with a model-based empirical analysis that examines liberalizing vs. non-liberalizing countries in the wake of the 1990 GATT Uruguay round\(^1\). Their results indicate that reductions in general tariffs are only weakly correlated with changes in GDP growth. However, reductions in tariffs on imported capital and intermediate goods have the most positive effect on GDP growth (Estevadeordal et al., 2007). Amiti and Konings’ (2007) work in Indonesia is consistent with this finding. Using Indonesian manufacturing data from 1991 to 2001 the authors find that reducing input tariffs increases overall productivity by approximately twice as much as decreasing output tariffs (Amrati et al., 2007).

These findings suggest that while reductions in tariffs may induce gains from trade, they are most effective when targeted at specific parts of the production process, namely intermediate inputs. But how do tariff reductions fare when compared to other trade barrier reductions, such as decreasing transportation costs? Following Melitz (2003), Balistrieri, et al. (2011) conduct econometric calibrations of a general equilibrium trade

\(^1\)Note that the authors use a 'treatment' and 'control' framework, where 'treatment' represents countries that liberalized after 1990 and 'control' represents those that had either already liberalized prior to 1990 or did not liberalize at all.
model to analyze the impact of reductions in tariffs and trade costs on welfare. The results indicate that while tariff reductions create welfare gains four times larger than the baseline, declines in fixed trade costs increase welfare by more than twice that of tariff reductions (Balistrieri et al., 2011). When the two are combined, the effect on welfare gains is even higher.

The Habsburg case speaks well to this literature, as it will allow me to test between the importance of the administrative reductions in trade barriers versus physical reductions. Habsburg trade liberalization was characterized both by reductions in tariffs via bilateral and multilateral trade agreements, as well as declines in trade costs through the massive construction of railroads. The section, below, provides an overview of this context.

3 Motivation – The Habsburg Natural Experiment

A review of the economic history literature reveals that the economic development patterns of the 19th century Habsburg Empire provide a relevant natural experiment in which to explore the above perspectives regarding trade and structural change. By the early 19th century, The Kingdom of Hungary (Hungary, hereafter) resided under the rule of the Austrian Habsburg Empire, but retained moderate political autonomy, with representation of its own diet in the imperial parliament. While trade had remained virtually free amongst the sub-regions of the Habsburg lands since the introduction of reforms by Maria Theresa in the late 18th century, tariffs between Austria and Hungary persisted well into the mid-19th century (Komlos, 1983). That is, tariffs within Austrian provinces (e.g., Bohemia, Moravia, Galicia, Tirol, Krakau, etc.) had been virtually eliminated, while tariffs between these places and Hungary endured.

Many authors claim that a significant motivation for the persistence of tariffs was a result of Austrian strong-arming against Hungarian nobles who refused to pay land taxes. (Eddie, 1977; Komlos, 1983; Good, 1975).
After defeat by the Austrians in the revolutions of 1848, Hungary was fully incorporated into the Austrian Empire, losing all of its administrative power and political representation. Furthermore, because the Austrian Constitution prohibited internal tariffs, the existing tariffs between the two regions were forced to dissolve (Eddie, 1977). Effectively, a customs union between the two formed. 1850 marks the official creation of the customs union, but the years just prior to 1850, until two decades afterward (i.e., roughly, 1848 to 1873), are commonly cited in the literature as a period of political and economic liberalization throughout the Empire (Eddie, 1977; Good, 1975; Komlos, 1983). Internally, these reforms were characterized by the liberation of the serfs in 1848, the introduction of two jointstock banks in 1853 and 1857, the establishment of a central bank and common currency, and the abolition of the guild system in 1859 (Eddie, 1967). Externally, the Habsburgs followed the trade liberalization trends of the day, engaging in "liberal treaties" with England, Germany and other economic powers of the region, which limited Austrian specific duties and ad valorem taxes (Eddie, 1977).

With the creation of the Dual Monarchy (i.e., the Austro-Hungarian Empire) in 1867, Hungary regained political autonomy within the Empire, and in fact, this autonomy was much stronger than that of the beginning of the century. Each region maintained its own parliament, but the Dual Monarchy shared a military, as well as joint fiscal and foreign policy. Although external protectionist sentiments began to emerge in this period throughout Austrian and Hungarian politics, it is important to note that the customs union persisted until the collapse of the Empire in 1913 (Eddie, 1977). The years following the creation of the Dual Monarchy are typically cited as the return to external protectionism within the Empire, most notably with the crash of the Vienna stock market in 1873 and reactions to tariff hikes in Germany in the early 1880s (Eddie, 1977). Consequently, the assumption of the present study is that one may consider Austria-Hungary a relatively closed economy in which internal regions underwent permanent trade liberalization. In other words, we can think of the Empire as a world economy
with two markets that move from relative autarky to relative openness. This perspective is confirmed to some extent by the economic history literature, which claims that internal tariffs between Austria and Hungary were lower than the external tariffs with the rest of the world (Eddie, 1977; Komlos, 1983). Although external trade did exist, the bulk of Habsburg trade appears to have occurred between Austria and Hungary (Komlos, 1983).

While the trade policy story is an important one, the current study is not able to empirically identify the effects of these changes, due to a dearth of data availability. Future work hopes to incorporate observations from periods before 1842 in order explore this issue further. The current study, instead, focuses on the effect of reductions in physical trade barriers, namely, railroad construction. 1837 marks the inauguration of the railway era, with the laying of 14 kilometers (km) of track in Austria. By 1870, rail construction began to boom with almost 2,000 km of track laid in that year alone. By the beginning of the 20th century, almost 45,000 km of track had been laid in the entire Dual Monarchy. Figures 1 and 2 illustrates the path of railroad construction between 1837 and 1914 in Austria and Hungary.

![Figure 1: Austro-Hungarian Railroad Construction: 1837-1910 - track length](image)

The economic history literature is not able to quantify the exact role railroad construction played in Habsburg economic development. However, numerous authors cite it as a potentially significant driver of growth through its linkages to other industries,
Figure 2: Austro-Hungarian Railroad Construction: 1837-1910 - track density

namely iron (Eddie, 1977); as a force of economic integration (Good, 1984); and as a solution to inefficient overland routes between markets that previously lacked access to the Danube (Rosegger, 1996). The goal of the present study is not to identify the railroad’s role in driving growth, but rather to examine its position in promoting integration and catalyzing structural change. A preliminary examination of the data reveals that structural change was indeed occurring over the course of the 19th century and may be correlated with the rise of the railroads. Figures 3 - 5 illustrate the trends in production structure.

Figure 3: Changes in Hungarian Production Structure: 1830-1915

From figures 3 and 4, it appears that in both Hungarian agriculture and Austro-Hungarian manufacturing production trends were fairly flat until around 1870 — the beginning of the railroad boom. Specifically, in Hungarian agriculture, there is no dis-
tinction between output of wheat, rye or potatoes, until the late 1870s, in which we observe a stark increase in wheat and potatoes production, with virtually no change in the trend for rye production. While the gains in potato production are likely due to technological improvements in agriculture\(^3\), it is suspected that the rapid increase in wheat production is largely correlated with the increase in Hungarian flour production in the early 1870s. That is, the boom in flour production observed in the Hungarian data for the early 1870s may have created a demand for wheat inputs, thereby explaining the rapid increase in wheat production we observe in the late 1870s.

An examination of the Austrian data reveals that manufacturing trends were relatively flat between iron, cotton textiles and flour until approximately 1865, when textiles and iron began to emerge as the higher value-added goods. Although the major booms in imperial railroad construction did not occur until 1870, Austria had been undergoing steady construction since the 1840s (as opposed to Hungary, which did not experience any serious construction until 1855). Therefore, it is quite likely that railroad development would have affected the patterns of Austrian structural change as early as 1865.

\(^3\)The height of the Industrial Revolution in Eastern Europe occurred in the late-19th century. During this time, large improvements in agricultural technology were made, including the use of chemical fertilizers, crop rotation and reapers (Good, 1984; Eddie, 1967; Van Zanden, 1991; Bicki, 2001). Potatoes are a soil-intensive crop and would have presumably benefited greatly from these improvements in technology.
Nonetheless, Figure 4 reveals that there was a clear divergence away from flour production in favor of more intensive manufactured goods, such as cotton textiles and iron, suggesting strong patterns of structural change in Austria in the last half of the 19th century.

Figure 5 compares Austrian and Hungarian production, by manufactured good, and suggests evidence for relative specialization between the two regions. Because the figures are presented in gross value-added terms and noting that the Hungarian economy was significantly smaller than the Austrian economy, one should pay closer attention to slopes rather than vertical distances between trends when attempting to discern patterns of structural change. With this in mind, some clear patterns emerge: Austria developed a relative specialization in intensive manufactures, such as iron and cotton textiles, while Hungary specialized in lower intensity manufactures, such as flour. In addition, the emergence of these trends appears to be highly correlated with the introduction of railroads, as the trends begin to emerge around 1870.

The Austro-Hungarian narrative provides the motivation for the current study. That is, through the 19th century Habsburg economy, it appears that there may be empirical evidence to support the notion that reductions in transportation costs catalyze structural change. However, while the economic history narrates a story of core-periphery between Austria and Hungary, there is some evidence that within Austria, itself, similar patterns emerged (see Figure 6). Due to a lack of Hungarian data, the current analysis will focus solely on the specialization that emerged within Austrian provinces.

4 Theoretical Model

I have constructed a formal (math) 2x2 Ricardian trade model that provides the intuition for the empirical findings. However, since the current draft is largely concerned with the empirical identification, I do not work through the math in this section and, instead,
describe the main assumptions and predictions of the model.

The classic Ricardian model assumes that there are two regions, $i$ and $j$, and two goods, $x$ and $y$. The production of each good utilizes one factor, labor, with Ricardian technology, $a$. Technology is specific to each region and good, while labor is immobile between regions, but mobile between sectors. Consumers and producers face "iceberg" transportation costs such that only a fraction of each good shipped from region $i$ to $j$ arrives at the destination.

In the exposition of the model, I walk through 3 different equilibria: 1) Autarky, in which neither region trades with one another; 2) Partial Specialization, in which one region completely specializes in good $y$, while the other region continues to produce both $x$ and $y$; and 3) Complete Specialization. Each equilibrium reveals a threshold level of transportation costs that is needed to uphold the optimum. When the thresholds are examined together, we find that as transportation costs decrease, the optimal special-
Figure 6: Production Specialization in Austrian Provinces: 1842-1912

The model, thus, predicts that as transportation costs decline, each region specializes in the good for which it has comparative advantage.

To garner any potential welfare effects of specialization, I explore comparative statics on the effect of declining transportation costs on prices and wages within the regional economy. I limit the analysis to the partial equilibrium, since this most closely resembles the Habsburg case in which regions such as Galicia and Bukowina specialized solely in agricultural production, while Bohemia and Lower Austria produced both agricultural and industrial goods throughout the 19th and early 20th centuries. The comparative statics reveal that the Partial Specialization equilibrium is upheld when one region has absolute advantage. In this equilibrium, as transportation costs decline, prices between
region $i$ and $j$ should converge, while wages should increase, but diverge.

5  Empirical Strategy

5.1  Basic Identification

The theoretical model suggests three different empirical specifications: 1) a regression of specialization on transportation costs, 2) a regression of pairwise price differentials on transportation cost, and 3) a regression of pairwise wage differentials on transportation costs. The results, below, therefore estimate the following regressions:

$$
|l_{it} - l_{jt}| = \alpha + \beta \text{connect}_{it} + \gamma \text{connect}_{ijt} + \rho X_{it} + \sigma_i + \mu_t + \epsilon_{ijt}
$$

(1)

where, \(\text{connect}_{it}\) is the number of railroad connections within province $i$ in year $t$, \(\text{connect}_{ijt}\) is the number of direct railroad connections between province $i$ and $j$ in year $t$, $X_{it}$ is a vector of controls for region $i$, and $\sigma_i$ and $\mu_t$ represent dyad (i.e., a given $ij$ pair) and time fixed effects, respectively. Note that the dependent variable, $l_{it}$, is a particular measure of specialization, known as "localization", where: $l_{it} = \frac{y_{ik}^i}{\sum_j y_{jk}^i} = "localization"$ — region $i$'s production of a particular good as a proportion of the total output of that good within the Empire. Taking the absolute value differences of each region's localization in year $t$ captures the extent to which some regions are becoming more localized and others are not. The larger the difference in these metrics, the more specialized the regional economy becomes. Therefore, one would expect $\gamma$ to be positive (and possibly $\beta$, depending on the effect internal transportation networks have on specialization. Note, however, that the simple theoretical model assumes that internal transportation costs have no effect on specialization).

In addition, I estimate the following regressions:
\[ |p_{it} - p_{jt}| = \alpha + \beta \text{connect}_{it} + \gamma \text{connect}_{ijt} + \rho X_{it} + \sigma_i + \mu_t + \epsilon_{ijt} \]  \hspace{1cm} (2)

\[ |w_{it} - w_{jt}| = \alpha + \beta \text{connect}_{it} + \gamma \text{connect}_{ijt} + \rho X_{it} + \sigma_i + \mu_t + \epsilon_{ijt} \]  \hspace{1cm} (3)

where, \( p_{it} \) is the price of good \( x \) in region \( i \) in year \( t \), and \( w_{it} \) is real wages\(^4\). If the data fit the theoretical model, we would expect the coefficient on \( \gamma \) in equation (2) to be negative, while the coefficient on \( \gamma \) in equation (3) should be positive. That is, if prices are converging as transportation costs decrease, we would expect that the more connections regions \( i \) and \( j \) have to one another, the smaller their price differences would be. Similarly, if wages are diverging, more connections between \( i \) and \( j \) means greater wage differences over time.

An alternative specification for the above equations is to use a dyad regression approach (see Fafchamps and Gubert (2007) for a detailed exposition on this methodology). Since the dependent variables are constructed such that they represent a specific \( ij \) pair, it may be important to control for characteristics that are particular to the pair, and not just the originating region, \( i \). A common way of estimating these data are via dyad regressions. In this case, equations (1), (2), and (3) would become:

\[ |l_{it} - l_{jt}| = \alpha + \beta_1 |\text{connect}_{it} - \text{connect}_{jt}| + \beta_2 (\text{connect}_{it} + \text{connect}_{jt}) + \gamma \text{connect}_{ijt} + \rho X_{it} + \sigma_i + \mu_t + \epsilon_{ijt} \]  \hspace{1cm} (4a)

\[ |p_{it} - p_{jt}| = \alpha + \beta_1 |\text{connect}_{it} - \text{connect}_{jt}| + \beta_2 (\text{connect}_{it} + \text{connect}_{jt}) + \gamma \text{connect}_{ijt} + \rho X_{it} + \sigma_i + \mu_t + \epsilon_{ijt} \]  \hspace{1cm} (4b)

\[ |w_{it} - w_{jt}| = \alpha + \beta_1 |\text{connect}_{it} - \text{connect}_{jt}| + \beta_2 (\text{connect}_{it} + \text{connect}_{jt}) + \gamma \text{connect}_{ijt} + \rho X_{it} + \sigma_i + \mu_t + \epsilon_{ijt} \]  \hspace{1cm} (4c)

That is, it may be necessary to control for the joint effect of each region’s own railroad connections, in addition to the pairwise connection. The coefficient on the absolute value term represents the effect of differences in connections, while the coefficient on the sum

\(^4\)Real wages were calculated by dividing nominal wages in region \( i \), year \( t \), by the price of wheat in region \( i \), year \( t \).
term represents the level effect. We may expect $\beta_2$ to have the same sign as $\gamma$ (since more connections in both places suggest further decreasing transportation costs), but it is unclear what sign $\beta_1$ may be.

Note that with dyad regressions, the standard errors tend to be larger if they are not corrected (Fafchamps and Gubert, 2007). Fafchamps and Gubert (2007) have written a STATA program to correct for the standard errors in dyad regressions. However, note that it only functions properly when there are no missing observations in the dependent variable. In addition, it is unclear how to correct standard errors when the dyad variables have been instrumented (I discuss my instrumental variables strategy in more detail, below). Unfortunately, historical data for this region and time period is scant, such that there are a number of missing observations in my dependent variables. In addition, to control for simultaneity bias in the placement of railroad connections, I implement an instrumental variables approach to equations (1)-(4). Therefore, the dyad regression results presented, below, do not correct for the skewed standard errors. Future research seeks to find corrections for these biases. Consequently, the empirical results of equations (4a)-(4b) must be interpreted with caution.

5.2 Empirical Issues

A major issue with any study of massive infrastructure developments is that the placement of such projects is likely endogenously correlated with economic outcomes. That is, it seems intuitive that the Habsburgs would have elected to construct railroads in places that were experiencing industrial growth in order to move intermediates and final goods in and out of these regions more efficiently. Similarly, as railroads reached these places, industrial specialization would be catalyzed via the mechanisms outlined in the theoretical model, above. It will, therefore, be necessary to construct a proper counterfactual for the placement of these tracks. Donaldson (2008) addresses this issue in India by running a placebo estimation in areas that were earmarked for railroad
construction, but never saw the construction plans come into fruition. In the current study, I could run a similar placebo test using a set of railroad plans from 1864, which would be overlayed with a railroad map from 1913 in order to determine which lines from the 1864 plans were never built. I would then include a variable in equations (1)-(4) that is the density of unbuilt tracks in region \( i \). If unobservable determinants of localization and price and wage differences decided the placement of railroads, then we would expect to see correlation between the unbuilt lines variable and the dependent variables, when controlling for regional and time fixed effects. However, if the decline in transportation costs associated with railroad construction is truly the one-way driver of production specialization, then we would expect the coefficient on the unbuilt track variable to be statistically insignificant. While the placebo approach is appealing, it cannot completely refute the endogeneity bias. In addition, the data is not yet ready for such a specification. Therefore, the current project implements a 2SLS instrumental variables approach.

In searching for an acceptable instrument, there is little narrative (in English) on the motivations for railroad placement in the Habsburg Empire. Rosegger’s (1996) essay — while providing data solely for 19th century Transylvanian railroads — claim to represent an accurate narrative of the nature of railroad construction throughout the Empire (Rosegger, 1996). In his account, placement of rail lines was characterized by tensions between local versus Imperial interests. While the Empire desired to place lines that would result in better access to maritime waterways (e.g., Black Sea), local interests were often concerned with the density of tracks, rather than destinations (Rosegger, 1996). Eddie (1967) references the Austrian historian, Karl Bachinger, who claimed that railroad placement was motivated largely in the early stages by military strategies, as the seeds for the 1848 revolutions began to grow, and the desire to avoid linking regions with underlying separatist sentiments and ethnic conflict in the latter part of the century (i.e., post 1880) (Eddie, 1967). Consequently, the economic history literature seems to suggest
that the motivations for railroad placement were not entirely economically motivated. This notion lends itself to using percent of military population in region, \( i \), at year \( t \) as a potential instrument. However, one may argue in the case of military population that governments only place military presence in regions that are economically significant, thereby violating the exclusion restriction for military population as a good instrument for track density.

Banerjee, Duflo et al. (2012) implement a unique strategy for dealing with endogenously placed infrastructure networks in their analysis of the impact of railroad development on long-term economic growth in China. With highly disaggregated county-level data, the authors construct (Euclidean) straight lines between cities that had a direct railroad connection in a given year. Excluding terminal nodes, any city that was in the direct path of the straight line was an exogenous beneficiary of the newly constructed transportation network. The authors then construct their instrument as the distance of county \( i \)'s centroid from the (exogenously determined) straight line in year \( t \) as a proxy for transportation costs. While this is an empirically sound method for introducing exogeneity into infrastructure regressions, my data is not disaggregated enough to implement such an approach.

Consequently, I implement an alternative set of instruments that exploit spatial and temporal determinants of railroad construction. The Habsburg trade statistics show that in the middle periods (i.e., 1860s) of railroad construction, the Empire imported a significant number of railroad tracks. While the data does not provide the country of origin, other data relating to bilateral trade with Germany reveals a large trade deficit (i.e., an excess of imports) with Germany in regard to 'fuels and construction materials'. Meanwhile, the consensus in the historical literature seems to be that Germany financed a large portion of the construction of the lines. It would, therefore, seem that railroad construction decisions were spatially dependent on proximity to Germany.

In addition, according to historical documents, the demand for coal was a particular
motivator for railroad construction (Entwurf, 1864). Coal was not only a necessary input into railroads, as it powered locomotives, it was also vital to industrial production. The historical argument claims that the relationship between coal mining, iron production and railroad development was so intertwined that the development of the railroad was in part responsible for the dearth of coal in Austria, as it required large amounts of coal to operate, and at the same time, constituted a means through which this dearth of coal could be ameliorated, should the construction of a more comprehensive railroad network be achieved. Therefore, it would seem that railroad construction decisions were temporally dependent on world coal prices. It is important to note that while world coal prices may affect industrial production, they should only have an effect on levels. That is, assuming that the production function for industrial goods is the same across regions, an exogenous fluctuation in the world price of coal would affect production functions equally across regions. This should have no consequent effect on regional industrial specialization. Instead specialization patterns should remain static and only levels of production should be affected.

Therefore, a proposed instrument is to interact distance from Munich – the closest German industrial center to the Austro-Hungarian Empire – with world coal prices to instrument for track density. The first stage regression takes the following form:

\[ \text{connect}_{it} = \alpha + \beta \text{DistMunich}_i \times \text{PriceCoal}_t + \rho X_{it} + \sigma_i + \mu_t + \epsilon_{ijt} \]  \hspace{1cm} (5b)

\[ \text{connect}_{ijt} = \alpha + \beta |IV_{it} - IV_{jt}| + (IV_{it} + IV_{jt}) + \rho_1 |X_{it} - X_{jt}| + \rho_2 (X_{it} + X_{jt}) + \sigma_i + \mu_t + \epsilon_{ijt} \]  \hspace{1cm} (5c)

where, \( IV_{it} = \text{DistMunich}_i \times \text{PriceCoal}_t \). Note that for the pairwise connections between province \( i \) and \( j \), I use a dyad regression identification in the first stage.

The estimates, below, calculate the distance to Munich as the shortest current auto route from province \( i \)'s capital to Munich. However, future estimations will use the
Euclidean distance from the centroid of the province. In addition, since I currently do not have data on world coal prices, the estimates below use Great Britain iron prices as a close proxy. Since iron was also an important input into railroad construction, the use of this proxy should not compromise the validity of the estimation.

6 Data Sources

To the great fortune of those seeking to study the Austro-Hungarian Empire, the Habsburgs were very thorough data gatherers. In particular, statistical yearbooks from 1841 to 1917 exist for both Austria and Hungary. While there is district-level disaggregation for some years, the majority of the statistics are at the provincial level. The statistical yearbooks provide a vast source of data, but it is important to note that, to date, none of this data has been digitized. At best, data exists in scanned pdf versions of the 19th century books; and at worst, they are stored in volumes of microfiche. Currently, I am in the process digitizing the yearbooks into a comprehensive electronic database. To date, I have entered data from 1842-1912 for each of the 14 Austrian provinces. Therefore, the results below exclude Hungary, Transylvania, and the military territories of Croatia-Slavonia, as well as the regions that are now parts of modern day Bosnia-Herzegovina and Serbia. Nonetheless, the yearbooks contain detailed information on agricultural output, prices, wages, industrial output, and other demographic variables, such as population, primary school enrollment, births, and deaths, among other socio-economic data. Much of these data were used to construct my dependent and control variables.

To define transportation costs, I gathered data on the sequencing of railroad construction throughout the Empire from historical statistical yearbooks, as well as railroad maps. The data from the railroad statistical books identifies the year in which a con-

---

5The current study uses following sources: Tafeln zur Statistik der österreichischen Monarchie (1842-1871), Östreichisches statistisches Handbuch für die im Reichsrathe vertretenen Konigreich und Lander (1882-1917), Ungarische Statistische Mitteilungen (1872-1910), and Ungarisches Statistisches Jahrbuch (1893-1913).
nection between two cities occurred. Using GIS, I then plotted each city on a 1913 railroad map and overlayed provincial administrative boundaries, in order to obtain the province(s) in which the connection occurred. I then drew straight lines between each city in a given connection in order to ascertain whether the connection was within or between provinces. This data was then used to construct dummy variables for years in which a railroad connection between province \( i \) and \( j \) opened (as well as years in which a connection within province \( i \) occurred). The \textit{connect} variables in the equations, thus, represent the aggregate number of connections within and between provinces in a given year.

Note that the railroad connections data in the following analysis is from 1842 to 1890, and for Austrian provinces, only. Hence, the estimations, below, are for 14 provinces, over approximately 50 years (i.e., 1842-1890). Future drafts will be able to incorporate data up to 1907 and will also include the seven Hungarian provinces.

7 Results

Table 1 presents summary statistics for the key dependent and control variables. Note that each of the dependent variables represents the absolute value difference between pairwise \( ij \) observations. In addition, note that all prices are in nominal Kronen (the late 19th century currency of the Empire), while wages are in real Kronen. The table reveals a fair degree of variation in the means of prices, with the largest differences arising in oat prices. Similarly, the localization measures show significant variation in metrics, with sugarbeet production representing the largest differences. Note, however, that due to the small sample size, the regressions, below, do not estimate results for sugarbeets. Once additional data is collected, however, there may be a enough observations to so.

In terms of railroad connections, the average number of aggregate connections within a given province over the study period is approximately eight. However, note that there
Table 1. Summary Statistics: Variables of interest – 1842-1890

<table>
<thead>
<tr>
<th>Dependent Variables¹</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheat Prices</td>
<td>3.82</td>
<td>3.31</td>
<td>0</td>
<td>21.75</td>
<td>9,214</td>
</tr>
<tr>
<td>Rye Prices</td>
<td>3.42</td>
<td>2.91</td>
<td>0</td>
<td>20.84</td>
<td>9,227</td>
</tr>
<tr>
<td>Oats Prices</td>
<td>4.49</td>
<td>4.17</td>
<td>0</td>
<td>29.34</td>
<td>9,280</td>
</tr>
<tr>
<td>Barley Prices</td>
<td>4.15</td>
<td>3.39</td>
<td>0</td>
<td>23.52</td>
<td>8,958</td>
</tr>
<tr>
<td>Potatoes Prices</td>
<td>2.96</td>
<td>2.71</td>
<td>0</td>
<td>18.14</td>
<td>9,224</td>
</tr>
<tr>
<td>Real Wages (with board)</td>
<td>0.02</td>
<td>0.02</td>
<td>0</td>
<td>0.09</td>
<td>8,880</td>
</tr>
<tr>
<td>Wheat Localization</td>
<td>0.09</td>
<td>0.10</td>
<td>0</td>
<td>0.43</td>
<td>4,284</td>
</tr>
<tr>
<td>Rye Localization</td>
<td>0.10</td>
<td>0.11</td>
<td>0</td>
<td>0.59</td>
<td>3,838</td>
</tr>
<tr>
<td>Oats Localization</td>
<td>0.10</td>
<td>0.09</td>
<td>0</td>
<td>0.51</td>
<td>4,200</td>
</tr>
<tr>
<td>Potatoes Localization</td>
<td>0.11</td>
<td>0.14</td>
<td>0</td>
<td>0.61</td>
<td>4,250</td>
</tr>
<tr>
<td>Raw Iron Localization</td>
<td>0.12</td>
<td>0.12</td>
<td>0</td>
<td>0.54</td>
<td>3,115</td>
</tr>
<tr>
<td>Poured Iron Localization</td>
<td>0.13</td>
<td>0.15</td>
<td>0</td>
<td>0.62</td>
<td>3,197</td>
</tr>
<tr>
<td>Black Coal Localization</td>
<td>0.18</td>
<td>0.17</td>
<td>0</td>
<td>0.60</td>
<td>1,816</td>
</tr>
<tr>
<td>Brown Coal Localization</td>
<td>0.14</td>
<td>0.23</td>
<td>0</td>
<td>0.79</td>
<td>3,887</td>
</tr>
<tr>
<td>Sugarbeet Localization</td>
<td>0.28</td>
<td>0.23</td>
<td>0</td>
<td>0.71</td>
<td>592</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Controls</th>
<th>Mean</th>
<th>Std. Dev.</th>
<th>Min</th>
<th>Max</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own Connections</td>
<td>7.71</td>
<td>15.96</td>
<td>0</td>
<td>107.00</td>
<td>10,290</td>
</tr>
<tr>
<td>Connections from (i) to (j)</td>
<td>0.08</td>
<td>0.34</td>
<td>0</td>
<td>4.00</td>
<td>10,290</td>
</tr>
<tr>
<td>Population</td>
<td>1,499,572</td>
<td>1,667,455</td>
<td>145,435</td>
<td>6,607,816</td>
<td>9,780</td>
</tr>
<tr>
<td>School Enrollment</td>
<td>0.73</td>
<td>0.28</td>
<td>0.12</td>
<td>1.00</td>
<td>4,395</td>
</tr>
<tr>
<td>Distance to Munich (miles)</td>
<td>533</td>
<td>354</td>
<td>145</td>
<td>1,365</td>
<td>10,290</td>
</tr>
</tbody>
</table>

¹All dependent variables represent the absolute value of the difference between regions \(i\) and \(j\).

is a large degree of heterogeneity across the provinces. Similarly, the average number of direct connections between provinces is quite low – just under one in aggregate; although, with approximately 20 more years of data pending, this metric is likely to increase.

Figures 7-9 illustrate trends in a subset of the dependent variables over the course of the study period. From these trend lines it appears that there is some evidence in support of the main hypotheses of the study: 1) specialization is occurring throughout the Empire, namely in regard to industrial goods, such as iron production. (There is
less clear evidence in terms of agricultural production; however, this may support the assumption that the Empire was only partially specialized); 2) prices are converging, and 3) (real) wages are dispersing.

Figure 7: Localization Trends: 1842-1912

Figure 8 illustrates trends in (nominal) agricultural prices from 1829-1913. From the trend lines there appears to be strong evidence in favor of the hypotheses that prices converged throughout the Empire over the study period. Similarly, Figure 9 suggests that (agricultural) real wages (including board - i.e., a daily meal) were diverging.

With the empirical motivation of the summary statistics in mind, I conduct estimations of the equations outlined in section 5. Table 2 presents the results of the first stage regression, in which I attempt to construct an instrument for endogenously placed
railroads by using province i’s distance from Munich, interacted with world iron prices in year t, to predict railroad construction. Column (1) is a rough estimate of own connections (i.e., connections within province i), which indicates statistical support for the
IV, but with an unintuitive sign. As regional and year dummies are added in columns (2) and (3), however, the IV remains statistically significant, and with the expected sign (i.e., as distance to Munich and world iron prices increase, railroad construction should decrease). Columns (4) and (5) present the results for the dyad estimations for the pairwise connections. Noting that the standard errors have not been corrected, there may be support for the IV, but only when accounting for dyad and year fixed effects. In addition, there only appears to be a level effect, but the sign is in the expected direction.

Tables 3a-3c present the results of the second-stage estimation for prices and reveal fairly strong support for price convergence. Note that in each of the regressions, column (1) is a rough OLS estimation, column (2) instruments only for own connections, column (3) instruments for both own and pairwise connections, but does not use a dyad specification, and column (4) instruments for own and pairwise connections using a dyad regression in the first and second stages. Columns (3) and (4), particularly, indicate that price convergence occurred throughout the study period, with negative and statistically significant signs on the connection variables. The two exceptions are potatoes and barley, where the coefficients on own connections are positive and significant, suggesting that as region \(i\) gains more connections within province, prices diverge with region \(j\). This trend is perhaps due to a lack of specialization in the production of these goods. The localization estimates in Tables 5a-5c should shed further light on this issue.

Table 4 presents the results for the second stage estimation for real wages and reveals strong support for wage dispersion, with positive and significant coefficients on the connection variables in columns (3) and (4). The results suggest that an additional pairwise railroad connection between province \(i\) and \(j\) results in a \(0.003\) to \(0.01\) Kronen increase in (absolute value) wage differences. From a mean difference of \(0.02\) Kronen, this is a fairly significant effect.

Tables 5-6 reveal much more ambiguous effects on localization. In regard to agricultural localization, Tables 5a and 5b suggest counterintuitive results. Recall that the
coefficients on the railroad variables should be positive if localization is occurring (i.e., as railroad networks grow, some regions become more localized while others become less localized. The net effect should be that the absolute value difference in localization metrics should increase as transportation costs fall). However, the pairwise connection coefficients, in particular, suggest the opposite – as railroad connections grow, differences in agricultural production are shrinking. There are a couple potential explanations
for this finding. One is that there is measurement error in the data. There is little discussion (in English) of the data collection techniques of the Habsburg statistical records. However, the consensus among economic historians is that the data likely suffer from

Table 3a. Second Stage: Absolute Value Differences in Agricultural Prices

<table>
<thead>
<tr>
<th>Connections</th>
<th>Connections</th>
<th>Connections</th>
<th>Connections</th>
<th>Connections</th>
<th>Connections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wheat</td>
<td>Wheat</td>
<td>Wheat</td>
<td>Wheat</td>
<td>Wheat</td>
<td>Wheat</td>
</tr>
<tr>
<td>Own Connections</td>
<td>0.00751** (0.00332)</td>
<td>-0.0124*** (0.00492)</td>
<td>-0.00784** (0.00381)</td>
<td>0.0412** (0.00558)</td>
<td>-0.0156** (0.00552)</td>
</tr>
<tr>
<td>Connections from /to/</td>
<td>0.478*** (0.142)</td>
<td>-1.209*** (0.110)</td>
<td>-1.717*** (0.0969)</td>
<td>-1.817*** (0.0995)</td>
<td>0.250 (0.163)</td>
</tr>
<tr>
<td>Identification</td>
<td>2.30e-06* (1.30e-06)</td>
<td>-0.0146*** (0.00250)</td>
<td>0.0412*** (0.00324)</td>
<td>-0.0176*** (0.00324)</td>
<td>-0.0176*** (0.00324)</td>
</tr>
<tr>
<td>Population</td>
<td>-2.63e-07* (1.43e-07)</td>
<td>-2.56e-07*** (3.66e-08)</td>
<td>-3.51e-07*** (4.57e-08)</td>
<td>-3.43e-07*** (2.33e-08)</td>
<td>1.16e-07 (1.64e-07)</td>
</tr>
<tr>
<td>School Enrollment</td>
<td>-1.385*** (0.248)</td>
<td>-1.086*** (0.238)</td>
<td>-2.051*** (0.280)</td>
<td>-0.852*** (0.209)</td>
<td>0.0906 (0.286)</td>
</tr>
<tr>
<td>Constant</td>
<td>3.074*** (0.590)</td>
<td>7.087*** (0.214)</td>
<td>8.369*** (0.263)</td>
<td>6.823*** (0.151)</td>
<td>1.559** (0.677)</td>
</tr>
</tbody>
</table>

Identification: OLS  
Observations: 3,974
R-squared: 0.513

Table 3b. Second Stage: Absolute Value Differences in Agricultural Prices

<table>
<thead>
<tr>
<th>Identification</th>
<th>Identification</th>
<th>Identification</th>
<th>Identification</th>
<th>Identification</th>
<th>Identification</th>
</tr>
</thead>
<tbody>
<tr>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
<td>OLS</td>
</tr>
<tr>
<td>Identification</td>
<td>4.20e-05*** (2.54e-06)</td>
<td>-0.0157*** (0.00489)</td>
<td>0.0513 (0.0059)</td>
<td>0.0412*** (0.00324)</td>
<td>-0.0176*** (0.00324)</td>
</tr>
<tr>
<td>Population</td>
<td>-5.76e-07* (2.80e-07)</td>
<td>1.32e-07 (8.24e-08)</td>
<td>2.39e-07*** (8.85e-08)</td>
<td>-2.07e-08 (4.56e-08)</td>
<td>3.24e-07* (1.66e-07)</td>
</tr>
<tr>
<td>School Enrollment</td>
<td>8.191*** (0.486)</td>
<td>6.882*** (0.535)</td>
<td>7.400*** (0.541)</td>
<td>-2.165*** (0.410)</td>
<td>-0.398 (0.303)</td>
</tr>
<tr>
<td>Constant</td>
<td>6.309*** (1.157)</td>
<td>-0.378 (0.505)</td>
<td>-0.308 (0.509)</td>
<td>5.333*** (0.295)</td>
<td>1.423** (0.687)</td>
</tr>
</tbody>
</table>

Identification: OLS  
Observations: 3,974
R-squared: 0.513

Standard errors in parentheses  
*** p<0.01, ** p<0.05, * p<0.1
a fair degree of measurement error. Considering the difficulty involved in aggregating data from individual producers to the provincial level, particularly in an era that lacked precise technology, it would seem reasonable to expect a bit of measurement error.

The alternative explanation is that perhaps these estimates provide evidence of partial specialization. For instance, Figure 7 reveals slightly ambiguous trends in agricultural localization. The historically industrial region of Bohemia exhibits large shares of wheat and rye production throughout the study period, and even appears to be converg-
<table>
<thead>
<tr>
<th>Own Connections</th>
<th>(1) Real Wages</th>
<th>(2) Real Wages</th>
<th>(3) Real Wages</th>
<th>(4) Real Wages</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.000191***</td>
<td>0.000364***</td>
<td>0.000338***</td>
<td>0.0102***</td>
</tr>
<tr>
<td></td>
<td>(2.23e-05)</td>
<td>(2.94e-05)</td>
<td>(3.20e-05)</td>
<td></td>
</tr>
<tr>
<td>Connections from $i$ to $j$</td>
<td>-0.00267***</td>
<td>0.00252***</td>
<td>0.000628</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000952)</td>
<td>(0.000769)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$</td>
<td>Connections_i - Connections_j</td>
<td>-4.16e-05***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(1.57e-05)</td>
</tr>
<tr>
<td>Connections, $+$ Connections,</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.45e-07***</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(8.06e-09)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>-1.41e-09</td>
<td>-2.79e-09***</td>
<td>-1.11e-09***</td>
<td>-8.21e-10***</td>
</tr>
<tr>
<td></td>
<td>(9.60e-10)</td>
<td>(2.57e-09)</td>
<td>(2.99e-10)</td>
<td>(1.46e-10)</td>
</tr>
<tr>
<td>School Enrollment</td>
<td>0.0229***</td>
<td>0.00550***</td>
<td>0.0215***</td>
<td>-0.00257**</td>
</tr>
<tr>
<td></td>
<td>(0.00170)</td>
<td>(0.00167)</td>
<td>(0.00185)</td>
<td>(0.00130)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.0154***</td>
<td>0.0207***</td>
<td>0.00455***</td>
<td>0.0136***</td>
</tr>
<tr>
<td></td>
<td>(0.00401)</td>
<td>(0.00146)</td>
<td>(0.00174)</td>
<td>(0.000930)</td>
</tr>
</tbody>
</table>

**Identification**

<table>
<thead>
<tr>
<th>Identification</th>
<th>OLS</th>
<th>IV: Own Connections</th>
<th>IV: Own &amp; Joint Connections</th>
<th>IV: Own &amp; Joint Connections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyad and Year Effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>3,852</td>
<td>3,490</td>
<td>3,327</td>
<td>3,327</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.447</td>
<td>0.419</td>
<td>0.419</td>
<td>0.467</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

ing on production with the historically agricultural region of Galicia. As the model suggests, perhaps the overall trends in the Empire were more akin to partial specialization in agricultural production, such that the negative coefficients capture the convergence of certain regions, like Bohemia, that also produced industrial goods.

The estimates of industrial localization provide much stronger evidence of specialization. Tables 6a and 6b present the results. In most of the specifications, the coefficients on pairwise connections are positive and significant, indicating that as transportation costs decrease, localization increases in certain regions and decreases in others, such that
the differences in localization increase on the net over time. While the level effects in some of the dyad regressions are negative, they are quite close to zero, perhaps suggesting a fairly precise zero level effect of own connections. Nonetheless, the estimates suggest

30
that an additional railroad connection between regions $i$ and $j$ may lead to increases in localization differences anywhere from 1 to 6 percentage points for iron, and 11 to 20 percentage points for coal. With means ranging from 12 to 18 percentage points, these effects appear to be quite large. It will be interesting to see how the predictions change once more data is added.

### Table 6a. Second Stage: Absolute Value Differences in Industrial Localization

<table>
<thead>
<tr>
<th></th>
<th>(1) Raw Iron</th>
<th>(2) Raw Iron</th>
<th>(3) Raw Iron</th>
<th>(4) Raw Iron</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own Connections</td>
<td>0.000150</td>
<td>(0.000150)</td>
<td>0.000217</td>
<td>(0.000217)</td>
</tr>
<tr>
<td>Connections from $i$ to $j$</td>
<td>0.000656</td>
<td>(0.000656)</td>
<td>0.000483</td>
<td>(0.000483)</td>
</tr>
<tr>
<td>Connections - Connections</td>
<td>0.000302</td>
<td>(0.000302)</td>
<td>0.000109</td>
<td>(0.000109)</td>
</tr>
<tr>
<td>Connections + Connections</td>
<td>-5.84e-07</td>
<td>(8.68e-08)</td>
<td>-7.91e-07</td>
<td>(1.26e-07)</td>
</tr>
<tr>
<td>Population</td>
<td>1.08e-05</td>
<td>(1.08e-06)</td>
<td>0.00130**</td>
<td>(0.000152)</td>
</tr>
<tr>
<td>School Enrollment</td>
<td>-0.0124</td>
<td>(0.00814)</td>
<td>0.00429***</td>
<td>(0.000197)</td>
</tr>
<tr>
<td>Constant</td>
<td>-0.00199</td>
<td>(0.00750)</td>
<td>-5.81e-09**</td>
<td>(1.05e-07)</td>
</tr>
</tbody>
</table>

Identification

- OLS
- IV: Own Connections
- IV: Own & Joint Connections
- IV: OLS
- IV: Own Connections
- IV: Own & Joint Connections
- IV: Own & Joint Connections

Dyad and Year Effects

- YES
- YES
- YES
- YES
- YES
- YES
- YES

Observations

- 1,659
- 1,472
- 1,428
- 1,428
- 1,794
- 1,635
- 1,592
- 1,592

R-squared

- 0.714
- 0.663
- 0.663
- 0.655
- 0.748
- 0.671
- 0.669
- 0.515

---

8 Conclusion

The above estimations reveal a fair degree of support for the main predictions of the model: 1) Specialization is occurring, albeit namely in industrial goods; 2) prices are converging, and 3) wages are diverging. Noting that the results are driven by a preliminary subset of the overall data, the estimations should be read cautiously. As more data is added for the seven Hungarian provinces and over an additional 20 years (i.e., up to 1907), it will be interesting to see how the empirical findings develop.

Future econometric work should focus on correcting the standard errors in IV dyad
### Table 6b. Second Stage: Absolute Value Differences in Industrial Localization

<table>
<thead>
<tr>
<th></th>
<th>(1) Black Coal</th>
<th>(2) Black Coal</th>
<th>(3) Black Coal</th>
<th>(4) Black Coal</th>
<th>(1) Brown Coal</th>
<th>(2) Brown Coal</th>
<th>(3) Brown Coal</th>
<th>(4) Brown Coal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Own Connections</td>
<td>-0.000202</td>
<td>-0.00306***</td>
<td>-0.00547***</td>
<td>0.00283***</td>
<td>0.00415***</td>
<td>0.00358***</td>
<td>0.00018</td>
<td>0.000159</td>
</tr>
<tr>
<td></td>
<td>(0.000156)</td>
<td>(0.000369)</td>
<td>(0.000630)</td>
<td>(0.000161)</td>
<td>(0.000159)</td>
<td>(0.000183)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Connections from i to j</td>
<td>-0.08696*</td>
<td>0.0774***</td>
<td>0.204***</td>
<td>0.165***</td>
<td>0.0335***</td>
<td>-0.0874***</td>
<td>0.107***</td>
<td>-0.101***</td>
</tr>
<tr>
<td></td>
<td>(0.00496)</td>
<td>(0.00797)</td>
<td>(0.0102)</td>
<td>(0.0113)</td>
<td>(0.00396)</td>
<td>(0.00392)</td>
<td>(0.00374)</td>
<td>(0.00390)</td>
</tr>
<tr>
<td>Connections - Connections</td>
<td>0.00300***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.000664)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Connections + Connections</td>
<td>-1.98e-06***</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(2.47e-07)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Population</td>
<td>-1.15e-08</td>
<td>2.19e-08***</td>
<td>2.17e-08***</td>
<td>1.11e-08***</td>
<td>-6.49e-08***</td>
<td>2.06e-08***</td>
<td>-1.08e-08***</td>
<td>-9.85e-09***</td>
</tr>
<tr>
<td></td>
<td>(7.09e-09)</td>
<td>(3.69e-09)</td>
<td>(4.44e-09)</td>
<td>(4.22e-09)</td>
<td>(7.84e-09)</td>
<td>(1.66e-09)</td>
<td>(1.96e-09)</td>
<td>(1.05e-09)</td>
</tr>
<tr>
<td>School Enrollment</td>
<td>0.0486*</td>
<td>0.319***</td>
<td>0.205***</td>
<td>0.220***</td>
<td>-0.00013</td>
<td>0.137***</td>
<td>0.301***</td>
<td>0.157***</td>
</tr>
<tr>
<td></td>
<td>(0.0257)</td>
<td>(0.0203)</td>
<td>(0.0255)</td>
<td>(0.0294)</td>
<td>(0.00976)</td>
<td>(0.0117)</td>
<td>(0.0136)</td>
<td>(0.0133)</td>
</tr>
<tr>
<td>Constant</td>
<td>0.0145</td>
<td>-0.0783***</td>
<td>-0.0374</td>
<td>-0.0179</td>
<td>0.0182</td>
<td>-0.0497***</td>
<td>-0.0767***</td>
<td>0.125***</td>
</tr>
<tr>
<td></td>
<td>(0.0302)</td>
<td>(0.0189)</td>
<td>(0.0229)</td>
<td>(0.0243)</td>
<td>(0.0190)</td>
<td>(0.0118)</td>
<td>(0.0135)</td>
<td>(0.00838)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Identification</th>
<th>OLS</th>
<th>IV: Own Connections</th>
<th>IV: Own &amp; Joint Connections</th>
<th>IV: Own Connections</th>
<th>IV: Own &amp; Joint Connections</th>
<th>OLS</th>
<th>IV: Own Connections</th>
<th>IV: Own &amp; Joint Connections</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dyad and Year Effects</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
<td>YES</td>
</tr>
<tr>
<td>Observations</td>
<td>1,059</td>
<td>799</td>
<td>799</td>
<td>799</td>
<td>2,247</td>
<td>2,287</td>
<td>2,255</td>
<td>2,255</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.917</td>
<td>0.731</td>
<td>0.649</td>
<td>0.520</td>
<td>0.966</td>
<td>0.966</td>
<td>0.894</td>
<td>0.898</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Regressions with missing observations and teasing out ambiguities in the partial specialization hypothesis. Nonetheless, the naive results appear to (at least weakly) uphold the theoretic predictions of the model.
References


