

Nitsch, Volker

**Working Paper**

## Openness and urban concentration in Europe, 1870-1990

HWWA Discussion Paper, No. 121

**Provided in Cooperation with:**

Hamburgisches Welt-Wirtschafts-Archiv (HWWA)

*Suggested Citation:* Nitsch, Volker (2001) : Openness and urban concentration in Europe, 1870-1990, HWWA Discussion Paper, No. 121, Hamburg Institute of International Economics (HWWA), Hamburg

This Version is available at:

<https://hdl.handle.net/10419/19430>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



DISCUSSION PAPER

# Openness and Urban Concentration in Europe, 1870–1990

**Volker Nitsch**

HWWA DISCUSSION PAPER

121

HWWA-Institut für Wirtschaftsforschung-Hamburg

2001

ISSN 1616-4814

The HWWA is a member of:

- Wissenschaftsgemeinschaft Gottfried Wilhelm Leibniz (WGL)
- Arbeitsgemeinschaft deutscher wirtschaftswissenschaftlicher Forschungsinstitute (ARGE)
- Association d'Instituts Européens de Conjoncture Economique (AIECE)

# **Openness and Urban Concentration in Europe, 1870–1990**

**Volker Nitsch**

This paper, prepared for the HWWA Conference on “Glocalization in Europe“, Hamburg, May 3-5, 2000, is based on chapter 4 of my dissertation “City Growth in Europe“ (Nitsch 2001a). I am grateful to my dissertation supervisors Dalia Marin and Michael Burda for their encouragement and very helpful discussions, and to Gianmarco Ottaviano, participants at the HWWA conference, and seminar participants at Humboldt University Berlin and the University of Munich (CES) for valuable comments. Opinions expressed in this paper do not necessarily represent the views of Bankgesellschaft Berlin.

## **HWWA DISCUSSION PAPER**

**Edited by the Department  
EUROPEAN INTEGRATION  
Head: Dr. Konrad Lammers**

Hamburgisches Welt-Wirtschafts-Archiv (HWWA)  
Hamburg Institute of International Economics  
Öffentlichkeitsarbeit  
Neuer Jungfernstieg 21 - 20347 Hamburg  
Telefon: 040/428 34 355  
Telefax: 040/428 34 451  
E-mail: [hwwa@hwwa.de](mailto:hwwa@hwwa.de)  
Internet: <http://www.hwwa.de/>

Volker Nitsch  
Bankgesellschaft Berlin  
Volkswirtschaft  
Hardenbergstrasse 32 - 10623 Berlin  
Tel.: (030) 3109-4462  
Fax: (030) 3109-5031  
E-Mail: [volker.nitsch@bankgesellschaft.de](mailto:volker.nitsch@bankgesellschaft.de)

# Contents

	<b>Page</b>
<b>1</b>	<b>INTRODUCTION</b> 7
<b>2</b>	<b>POTENTIAL CAUSES FOR URBAN CONCENTRATION</b> 9
2.1	Economic Development 9
2.2	Political Power 11
2.3	Transportation Infrastructure 11
2.4	Regression Framework 12
<b>3</b>	<b>DATA</b> 13
3.1	Data Sources 13
3.2	Alternative Measures of Urban Primacy 15
<b>4</b>	<b>RESULTS</b> 19
4.1	Replicating Ales and Glaeser (1995) 19
4.2	More Years of Data 27
4.3	Other Concentration Measures 30
4.4	Full Time Period, 1870-1990 32
4.5	Changes in Urban Concentration 33
<b>5</b>	<b>CONCLUSION</b> 37
	<b>References</b> 39
	<b>Appendix</b> 41
	<b>Tables</b>
Table 1:	Simple Correlations 17
Table 2:	Replicating Ales and Glaeser (1995) 20
Table 3:	Alternative Estimation Techniques and Sample Sizes 24
Table 4:	Additional Explanatory Variables 25
Table 5:	More Years of Data 28
Table 6:	Other Concentration Measures 31
Table 7:	Full Time Period, 1870-1990 33
Table 8:	Changes in Urban Concentration, 1870-1990 35
Table 9	Openness and Changes in Urban Concentration, 1870-1990 36
	<b>Figures</b>
Figure 1:	The Evolution of the Openness Ratio in Europe 8
Figure 2:	The Evolution of the Openness Ratio and Measures of Urban Concentration 18

## **Summary**

This paper provides a new attempt to examine the relationship between an economy's exposure to foreign trade and its degree of urban concentration. In contrast to previous work, which is often based on rich cross-country data sets, the analysis is focused exclusively on European countries, allowing to explore reliable historical data. The results generally confirm earlier findings of a negative relationship between openness and city size for the period from 1970 to 1990, but are much weaker for other time periods and concentration measures. Therefore, it is argued that the linkage between trade policy and urban centralization is not robust.

## **Zusammenfassung**

Dieses Papier untersucht den Zusammenhang zwischen dem Offenheitsgrad einer Volkswirtschaft und dem Ausmaß urbaner Konzentration. Im Unterschied zu früheren Studien, die oftmals auf umfangreichen Querschnittsdatensätzen basieren, beschränkt sich die Analyse in dieser Arbeit auf einige europäische Länder, für die dann allerdings verlässliche historische Daten verfügbar sind. Die Ergebnisse bestätigen frühere Resultate eines negativen Zusammenhangs zwischen dem Offenheitsgrad eines Landes und der Einwohnerzahl der größten Stadt für den Zeitraum von 1970 bis 1990. Für andere Zeiträume und Konzentrationsmaße verliert der Zusammenhang jedoch an Signifikanz, so dass es offenbar keine robuste Verbindung zwischen dem Grad der Außenhandelsverflechtung und dem Ausmaß urbaner Konzentration gibt.

**JEL Code:** R11, R12

**Keywords:** Urban Concentration, Openness, Urban Primacy, City Size

# 1 INTRODUCTION

One of the most striking empirical features in urban economics is the large difference in urban concentration across countries. In Austria, for example, about 20 percent of the total population live in the nation's largest city, Vienna, while in neighboring Switzerland, the comparable ratio is considerably smaller, with Zurich containing less than 6 percent of the population. As there is no obvious reason, why extremely different shares of a country's population are concentrated in the main city, a huge and still growing literature seeks to explain these differences in urban primacy.

An interesting recent contribution in the debate is a paper by *Krugman and Elizondo* (1996). Inspired by the giant size of Mexico City, they develop a simple theoretical model in which protectionist trade policies are a major cause of large central cities. More generally, arguing that "international trade theory and urban economics cannot, ultimately, be regarded as wholly separate disciplines" (p. 150), they suggest that there is a direct linkage between trade policy and urban concentration.

As this insight of *Krugman and Elizondo's* more or less theoretical exercise is, on the one hand, novel and fascinating and, on the other hand, quite easily testable, it is surprising that, to date, there has been relatively little effort to check this hypothesis empirically. In fact, the only thorough test I know of has been provided by *Ades and Glaeser* (1995) who find in a cross-section sample of 85 countries that the share of trade in GDP is indeed negatively related to the size of the largest city, providing mild support for *Krugman and Elizondo* (1996). Their results, however, are not very robust. The coefficient on openness, for example, loses significance if a Latin America dummy is included in the regressions. Moreover, *Ades and Glaeser* report some anecdotal evidence in the form of historical case studies where large cities **grew** as a result of trade and commerce, i.e., suggesting exactly the opposite relationship. In sum, the empirical evidence for *Krugman and Elizondo's* thesis that urban concentration is negatively related to international trade is both rare and far away from being convincing.

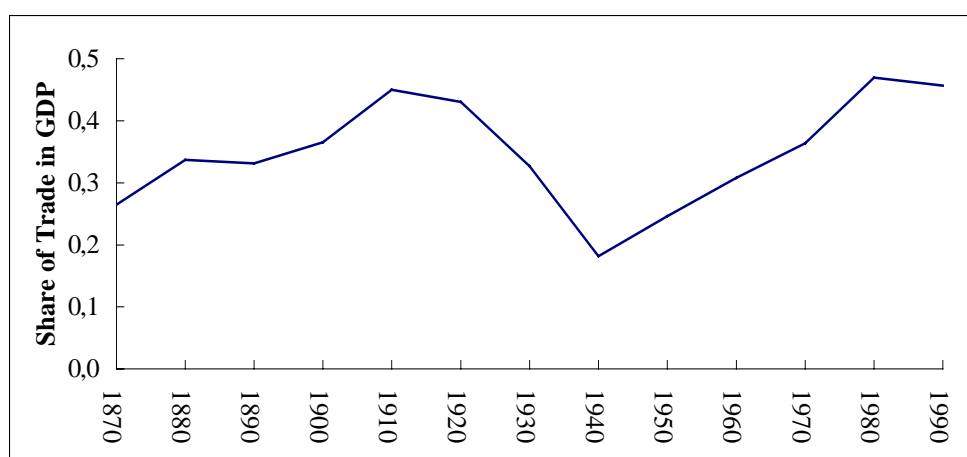
This paper provides a new attempt to examine the relationship between an economy's exposure to foreign trade and urban concentration. In particular, it contributes to the literature along several lines. First, the analysis is explicitly focused on only one region of the world, Europe. As a closer look reveals, this focus offers at least two advantages.



On the one hand, there is only a limited loss of information. In fact, European countries offer a wide variety of country and population sizes, population densities, per capita incomes, trade openness', and, as already noted, urban concentrations. On the other hand, if the analysis is focused on only a few countries, often a lot of additional information is available. In particular, the focus on Europe allows to examine (reliable) historical data.

Therefore, the second contribution of this paper is to analyze the association between trade policy and urban development not only in cross-country fashion but also in time series dimension. As figure 1 illustrates there has been considerable variation in the openness ratio of European countries across time. While the aggregated trade-to-GDP ratio has more than doubled in the time period from 1870 to 1913, it collapsed from 1913 to 1945 and then recovered to its pre-World War I level from 1945 to 1990. If *Krugman* and *Elizondo's* (1996) theoretical reasoning is correct, one should expect that such huge long-term shifts in the openness of European countries have had a considerable impact on the evolution of large central cities in Europe.

**Figure 1      The Evolution of the Openness Ratio in Europe**



Notes: The share of trade in GDP is calculated as the weighted average of the trade-to-GDP ratio for individual countries.

Finally, the empirical analysis in this paper is not confined to only one measure of urban concentration. While *Ades* and *Glaeser* (1995) use almost exclusively the absolute size

of a country's main city as their dependent variable, this paper examines the impact of trade policy on several measures of urban centralization in a country.

This paper, then, is in five parts. Section 2 describes the basic determinants of urban primacy and the regression framework. Section 3 gives a brief description of the data. Section 4 presents the results, and Section 5 concludes.

## **2 POTENTIAL CAUSES FOR URBAN CONCENTRATION**

Generally, the question about discrepancies in urban primacy around the world, i.e., why the size of the largest city relative to total population differs across nations, is only a variant of the more fundamental question why economic activity is not evenly distributed across space. As this problem is at the heart of location theory, one of the oldest fields in economics, a huge literature has evolved over the last century which aims to investigate the factors behind the wide variation in the size of urban metropolises.<sup>1</sup>

Not surprisingly, several explanations have been proposed in the literature. As it seems hard, however, to believe that large central cities are the result of only a single exogenous factor (e.g., a protectionist trade regime), this section explores a number of reasons which can plausibly contribute to an empirical explanation of the observed differences in urban concentration. In particular, the analysis focuses – besides trade policy – on three forces: (i) economic development; (ii) concentration of political power; and (iii) transportation infrastructure.

### **2.1 Economic Development**

To urban economists it is very well known that there are close interactions between urbanization and the economy. Examining the sign of the relationship between economic development and urban concentration, however, there are equally convincing

---

<sup>1</sup> The earliest study which *Carroll* (1982) cites in his interesting review of empirical studies about city size distributions is a German paper published in 1913.

arguments for both sides. On the one hand, it is quite obvious that if a large fraction of a country's population is working in sectors which depend on immobile natural resources such as agriculture and mining, the scope for urbanization will be limited. In the extreme case of an economy solely based on agriculture (e.g., in a society of hunters and gatherers), there will be no population concentration at all. This reasoning implies, however, that industrialization, i.e., a shift in economic structure away from the primary sector of an economy (often associated with economic development)<sup>2</sup>, will raise the level of urbanization and, accordingly, increase the extent to which an economy can centralize in a single location.

Moreover, agglomeration and urban concentration itself provide economic benefits (agglomeration economies). By concentrating in the same place, firms can benefit from a pooled labor market for specialized workers, a larger variety of inputs in production, and informational spillovers, as has been already noted in the late nineteenth century by *Marshall* (1890 [1920]). But if cities have attributes which positively affect the productivity and the growth of the economy, then urban centralization may also be positively correlated with economic development.<sup>3</sup>

On the other hand, there is also an intuitive line of reasoning suggesting exactly the opposite relationship. As an economy industrializes, this tends to increase the size of the domestic market. When local demand thresholds are passed, a growing number of firms will find it attractive to locate away from the center to serve regional markets in order to save transport costs. High income, then, should allow a country to support a network of intermediate-sized cities. In fact, *Rosen and Resnick* (1980) find that wealthier countries have more evenly distributed populations.

Given this ambiguity about the sign of the linkage, a number of economists have even argued that there is a nonmonotonic inverted U-shape relationship between economic development and urban concentration, i.e., that economic development initially increases and then decreases urban primacy. The empirical evidence on this point, however, has been mixed. While time series analyses often provide support for the inverted U-curve hypothesis, cross section evidence has been less conclusive.

---

2 *Ades and Glaeser* (1995), for example, report a sample correlation of 0.849 between the share of the labor force outside of agriculture and GDP per capita.

3 An alternative hypothesis would be that economic development is correlated with urbanization and population density, but not with urban concentration.

## 2.2 Political Power

It is one of the contributions of recent empirical work in the new economic geography literature to have reemphasized the role of government and politics in determining city size. *De Long* and *Shleifer* (1993), for example, present evidence showing that absolutist governments have inhibited city growth in preindustrial Europe. Discussing the failure of Zipf's law in many countries due to unproportionately large central cities, *Krugman* (1996b, p. 41) notes that "[t]hese primate cities are typically political capitals; it is easy to imagine that they are essentially different creatures from the rest of the urban sample." Most explicitly, however, *Ades* and *Glaeser* (1995, p. 195) argue that "political forces, **even more than economic factors**, drive urban centralization" (emphasis added).

The basic idea behind an association between politics and urban primacy is that spatial proximity to political power usually increases political influence. In some cases, then, there can be strong incentives to be near the government. Dictatorships, for example, often ignore the needs of the politically weak hinterland so that it may be highly attractive for rural workers to move to the capital. Also weak governments can be expected to raise the size of the central city: By transferring resources to the capital, they try to enhance the probability of their re-election and effectively attract rent-seeking migrants. In sum, undemocratic institutions, political instability and dictatorships should favor urban concentration.

## 2.3 Transportation Infrastructure

The bulk of urban and regional economics explains the demand for cities by the desire to minimize transport costs. When transportation is costly, firms can eliminate distance by locating close to each other. Thus, the incentive for industries to cluster together will, at least in part, depend on internal transport costs. The less efficient the national transport network and, thus, the more expensive the transportation of inputs and final products, the more firms will concentrate in one location. Alternatively, a fall in the costs of moving goods should lead to a more equal distribution of economic activity.

However, a better infrastructure does not always decrease urban concentration. In fact, there is an equally intuitive argument which suggests that an improvement of the

physical infrastructure of a country may also **promote** urban centralization. Only the ability to transport goods over longer distances made agglomerations possible in the first place (see *Bairoch* 1988). More generally, since lower transportation costs allow a city to service a larger hinterland, an efficient transport network may generate larger central cities. *Krugman* (1991) formalizes this effect. In his two-regions-model, a fall in transportation costs works in favor of regional divergence (through the possible realization of stronger backward and forward linkages) where, for transport costs below a critical level, complete concentration in one region is an equilibrium. Reviewing the empirical evidence, *Krugman* (1996, p. 13) refers to results from *Rosen* and *Resnick* (1980) and notes that "[c]ountries in which the capital city has a uniquely central position ... tend, not too surprisingly, to have more populous capitals".

Taken together, these points imply that the association between internal transport costs and concentration is ambiguous. Thus, by controlling for a country's transportation infrastructure, I implicitly test for both alternative hypotheses.

## 2.4 Regression Framework

In sum, I estimate a regression of the form:

$$(1) \quad URBCONC_i = \alpha + \beta OPEN_i + \gamma Z_i + \varepsilon_i,$$

where  $i$  denotes the country,  $URBCONC$  is the measure of urban concentration,  $OPEN$  is the share of trade in GDP,  $Z$  is a vector of other explanatory variables, and  $\varepsilon$  is a randomly distributed disturbance term. If *Krugman* and *Elizondo's* (1996) hypothesis is correct, I would expect that the estimated coefficient on the variable of interest, openness, is negative (i.e.,  $\beta < 0$ ). Otherwise, there would be no empirical support for their model.

## 3 DATA

### 3.1 Data Sources

My sample comprises data from 13 European countries<sup>4</sup>, covering the time period from 1870 to 1990 in 10-year-intervals. The countries were selected to provide the most complete set of historical data. Nonetheless, I lose several observations mostly in dealing with the countries' population structure.

In compiling the data, there were basically two sets of problems. A first issue refers to inconsistencies in the construction of the city population figures which were taken from national statistical yearbooks. In particular, the definitions of metropolitan areas are likely to differ across countries. Sometimes the numbers comprise the whole agglomeration while in some cases they refer only to the central city. Accordingly, the city population data are not directly comparable between countries. As there is no indication, however, that the data in particular countries are consistently biased by explicitly referring to wider agglomerations instead of cities, the impact on the empirical results should be limited.

A related problem is that as agglomerations often grow over time and merge with surrounding cities, the geographical areas referred to by national statistical agencies may differ across time. With the formation of Greater Berlin in 1920, for instance, reporting census data for former cities such as Charlottenburg was ceased, and – according to national statistics – Berlin's population more than doubled within five years, rising from about 1.7 million in 1915 to 3.8 million in 1920. On average, however, this effect is likely to cancel out in the time series dimension. With growing agglomerations, the true impact of the explanatory variables on city size will be first underestimated and then overestimated.

A second difficulty is caused by the numerous changes which have taken place in national boundaries. In principle, it is possible to adjust for those frequent border redrawings (and, thus, abrupt changes in country characteristics such as land area and

---

<sup>4</sup> The countries are Austria, Belgium, Denmark, Finland, France, (West) Germany, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, and Switzerland.

total population) with no **immediate** impact on the population of a country's largest city by including a persistence variable. In this paper, however, I basically deal with this problem by taking most of the historical data series (e.g., level of GDP, GDP per capita) from *Maddison* (1995) who already corrects for territorial change. In particular, *Maddison's* data refer exclusively to the present territory of the countries.

The flipside of having a consistent basis in terms of geographical area is that at least in some cases urban primacy is not related to actual, but corrected country data. This problem, however, should be of minor importance, largely for two reasons. First, with the exceptions of Austria and Germany, territorial changes in most countries in the sample affected only a small proportion of the total area so that the impact on the overall country data should be negligible. In Spain and Portugal, for example, there were no border changes at all.

Second, whenever feasible, explanatory variables are entered into the regressions not in levels but in relations or shares. The idea is that territorial changes should have had only a rather limited impact on variables such as a country's overall GDP per capita or on the share of labor force outside of agriculture.

To minimize the errors introduced by the basic strategy of using data referring to present boundaries, however, I depart from the time-consistency approach in compiling the data for the two countries in the sample with the most dramatic territorial changes, Austria and Germany. Specifically, I allow for a marked reduction in country size and, thus, a break in level data (such as area, GDP, and population) in 1920 and 1950, respectively. As *Maddison's* (1995) adjusted data are then only applicable for the time period after that date, pre-break level data are constructed from *Mitchell* (1992) who reports actual population figures. Given this population data, GDP numbers comparable to *Maddison's* estimates are calculated by multiplying population with *Maddison's* GDP per capita. Thus, I implicitly assume that the frontier change has not affected the country's GDP per capita.<sup>5</sup>

A second modification is chosen for German data. Due to the special geographical and political position of Berlin after World War II, I refrain from using Berlin as Germany's

---

5 This assumption is justified by *Maddison* (1995, table H-2) who confronts actual and adjusted estimates of GDP per capita for Germany and reports only minor differences. In 1870, for example, actual German GDP per capita (1821 \$) is about 95% of *Maddison's* estimate for GDP per capita adjusted to 1990 frontiers (1913 \$).

largest city after 1945 and focus instead on the largest city in the contiguous territory of former West Germany, Hamburg.<sup>6</sup>

Having explained the basic procedures to derive the data, an appendix provides a detailed description of the data sources.

### 3.2 Alternative Measures of Urban Primacy

Given a variety of primacy measures in the literature, it is one of the contributions of this paper to allow for different indices of urban concentration as dependent variables. Specifically, four measures of urban primacy are used to examine the relationship between trade openness and urban concentration, each having its own advantages and shortcomings and throwing light on slightly different aspects of population concentration.

The first measure, then, is the absolute number of inhabitants in a country's largest city as recently used by *Ades* and *Glaeser* (1995). Strictly speaking, this is not a direct measure of concentration since the size of a country's largest agglomeration should be strongly affected by the country's total population, with more populous countries also having larger cities. Not surprisingly, then, *Ades* and *Glaeser* (1995, table 3) report that in their sample the country's main city size is most highly correlated with the country's total population (with a correlation coefficient of 0.537). However, after correcting for this relationship by including a country's total population into the regression as an explanatory variable, the population in a country's main urban agglomeration is a feasible measure of population concentration.

In a slight modification of this regression specification, the share of the largest city in total urban population is used as dependent variable. This ratio has been widely used in the literature (see, for example, *Moomaw* and *Shatter* 1996 for a recent application). The main difficulty with this index (and also with the first measure) is, however, that it largely ignores the size distribution of cities below the largest agglomeration. To compensate at least partly for this deficiency, variants of the primacy index have been formulated, covering a larger number of cities. *Junius* (1999), for example, has also calculated the share of the two, three, and four largest cities in total population,

---

6 None of the results reported are qualitatively affected by this choice.



respectively. His results, however, were basically identical for those modified indices as dependent variables. In my sample, the correlation between the ratio of the largest city to a country's total population and the ratio of the two largest cities to total population is 0.979.

A third primacy measure relates the largest city to the population in the second largest metropolitan area. Focusing exclusively on the upper bound of the size distribution of cities, this ratio illustrates more directly the dominance of a country's largest agglomeration. Again, however, the measure provides no information about the size distribution of cities below the top two largest. Variants of this index, therefore, include a larger number of cities in the denominator. *Rosen* and *Resnick* (1980), for example, calculate ratios of the largest city to the sum of the top five and the top 50 cities, respectively. However, their modifications provide largely no new insights as there is a strong correlation of 0.922 between the two measures. Also due to lack of sufficient data, then, the empirical analysis in this paper applies the most simple formulation of this index, focusing on the ratio between the two largest cities of a country.

Finally<sup>7</sup>, following *Wheaton* and *Shishido* (1981) and others, an *H* index of concentration is computed which is defined as

$$(2) \quad H = \sum_{i=1}^n \left( \frac{CITY_i}{POP} \right)^2,$$

where  $CITY_i$  is the population of city  $i$ ,  $POP$  is the country's total urban population, and  $n$  is the number of cities included in the calculation of the index. In contrast to the previous measures, the basic advantage of this index is that it considers the size distribution of *all* cities within a system. As is familiar from the application of comparable distribution measures in other contexts (such as market segmentation), the *H* index can vary between 0 and 1, with large values of *H* indicating a higher population concentration. In the extreme case of  $H=1$ , the total population of a country is concentrated in a single city. A number of studies also emphasize the interesting economic insight of the reciprocal of *H*, measuring the number of equal-sized cities which would generate the urban population.

---

7 The list of primacy measures discussed above and used in the following analysis is far from being exclusive. For example, an also frequently used measure, which offers the advantage of considering the entire spectrum of city sizes, is the Pareto exponent. Insufficient data, however, prevented the inclusion of this concentration measure in the empirical analysis of this paper.

The main difficulty in calculating the  $H$  index is the definition of the cut-off point, i.e., how many cities should be included in the computation of  $H$ . Here it has become common procedure to include the largest agglomerations of a country that account for a given **fraction** of the total urban population. An alternative definition of including all cities above a fixed population size, say larger than 100,000, would distort the results for small countries. In this study, the cut-off is (limited by data availability) set at 60% of the total urban population in a country and, thus, somewhat lower than, for example, *Wheaton and Shishido's* (1981) cut-off of 70%.

Having explained the details of constructing different measures of population concentration, table 1 gives simple correlation coefficients. Interestingly, the interaction is weak in most cases, supporting the assumption that different measures also reflect different aspects of urban concentration. The only correlation which is significantly high is between the share of the largest city in total urban population and the  $H$  concentration measure. Given the large similarity in constructing both measures, this finding is not too surprising, but also suggests that the additional insight gained by incorporating the city structure below the largest city is rather limited in this sample.

**Table 1: Simple Correlations**

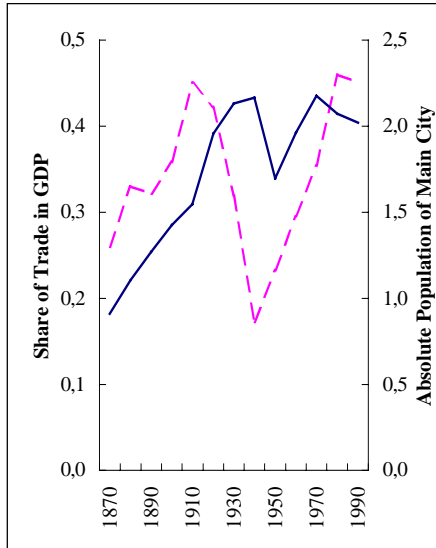
	MAINCITY	MAINSHARE	FIRSEC	HCONC
Absolute Population of Main City (MAINCITY)	1,00			
Share of Main City in Total Urban Population (MAINSHARE)	-0,20	1,00		
Ratio of Main City to Second Largest City (FIRSEC)	0,22	0,41	1,00	
H Concentration Measure (HCONC)	-0,15	0,94	0,34	1,00

Notes: All correlation coefficients are based on variable averages for the period 1970-90

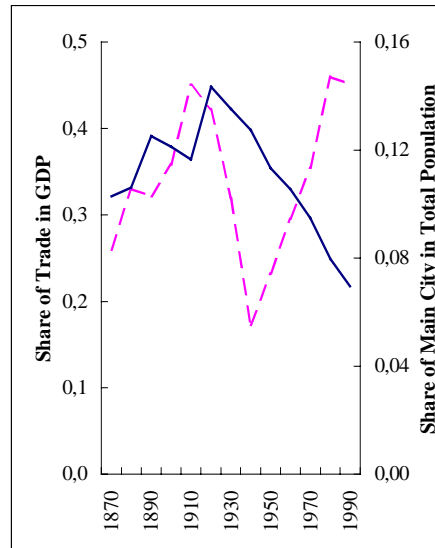
Figure 2 visualizes these observations showing weighted averages of the variables of interest across time. Given the familiar pattern in trade openness, no clear correlation with the various measures of urban concentration is evident. For example, the absolute size of the largest city has increased in almost every decade since 1870, apparently the

**Figure 2: The Evolution of the Openness Ratio and Measures of Urban Concentration**

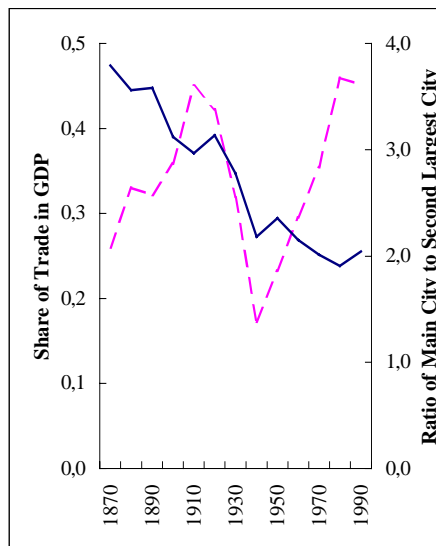
a) Absolute Population of Main City



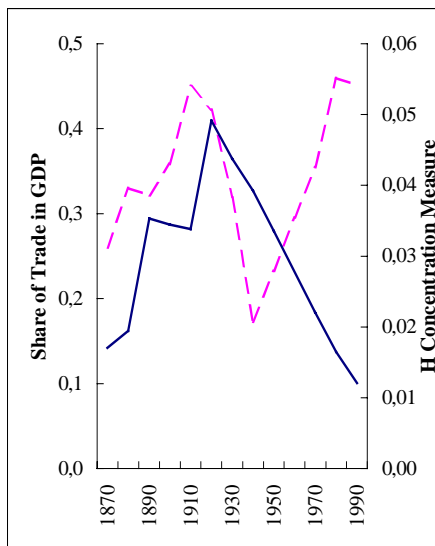
b) Share of Main City in total Population



c) Ratio of Main City to Second Largest City



d) H Concentration Measure



Notes: The solid line represents the concentration measure (right scale) and the dotted line is the openness ratio.

ratio between a country's largest and second largest city. A third pattern is provided by central city size corrected for a country's population. Both related measures, the share of the largest city in total population and the  $H$  concentration measure, have risen in the first half of the time period under investigation and have fallen since then to values even below the 1870 level.

## 4 RESULTS

### 4.1 Replicating Ades and Glaeser (1995)

The first set of estimates attempts to replicate the results in *Ades and Glaeser (1995)* who find in a cross sample of 85 countries a statistically significant negative relationship between the share of trade in GDP and the size of the largest city. Specifically, the aim is to analyze whether their finding is also valid for a European (sub)sample of 13 countries, which were – with the exception of Sweden – also all included in *Ades and Glaeser's* original analysis. Moreover, focusing on almost the same time period, averaging data for 1970, 1980 and 1990<sup>8</sup>, has the advantage of using exclusively high-quality data. Thus, possible problems in constructing the historical data set can be ignored. The basic shortcoming of this cross-country exercise, however, is the small sample size of only 13 observations which clearly limits the degrees of freedom. Robustness checks, therefore, apply alternative estimation methods which allow to explore jointly the information from separate years (instead of averaging data for the whole time period).

Table 2 presents the results. Column (1) repeats *Ades and Glaeser's (1995)* benchmark regression which relates the log of average population in the main city to a standard set of controls: a capital city dummy, the log of non-urbanized population, the log of urbanized population outside the main city, the log of land area, the log of real per capita GDP, and the share of the labor force outside of agriculture. None of these explanatory variables has a coefficient that is statistically significant at conventional levels of confidence. Given that also in *Ades and Glaeser's* study with 85 observations not all estimated coefficients are significant<sup>9</sup>, this is not too disappointing. Interestingly, however, even the scaling variables such as population or land area which control for

---

<sup>8</sup> *Ades and Glaeser (1995)* use averages of 1970, 1975, 1980 and 1985 observations.

<sup>9</sup> Only the log of non-urbanized population, the log of land area and the share of labor force outside of agriculture have statistically highly significant coefficients.

the size of a country have no predictive power for the population size of the country's largest city. Taking into account that the adjusted  $R^2$  is nevertheless quite respectable at 0.72, this result suggests that the regression specification apparently misses an important explanatory variable. It is also possible that the result suffers from the small sample size and the limited degrees of freedom.

**Table 2: Replicating Ades and Glaeser (1995)**

Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of Population in Main City	Avg. 1970-90	Avg. 1970-90	Avg. 1970-90	Avg. 1970-90	1970	1980	1990
Constant	-0.026 (8.345)	22.527** (4.940)	22.300** (4.102)	14.430* (4.833)	3.685 (5.408)	12.398* (5.020)	22.484* (9.230)
Capital City Dummy	0.703 (0.368)	0.448** (0.111)	0.450** (0.106)	0.445** (0.123)	0.742* (0.224)	0.474* (0.146)	0.297 (0.279)
Log of Non-Urbanized Population	0.460 (0.302)	0.510** (0.122)	0.489** (0.047)	0.509** (0.076)	0.536** (0.134)	0.488** (0.086)	0.530** (0.085)
Log of Urbanized Population Outside the Main City	0.187 (0.220)	-0.022 (0.119)					
Log of Land Area	0.111 (0.098)	-0.259* (0.091)	-0.257* (0.082)				
Log of Real GDP per Capita	0.276 (0.781)	-2.940** (0.726)	-2.872** (0.498)	-1.855* (0.579)	-0.188 (0.478)	-1.763* (0.591)	-3.334* (1.257)
Share of the Labor Force Outside of Agriculture	0.154 (4.224)	16.384* (4.363)	15.833** (2.164)	11.229** (2.187)	4.810* (1.657)	12.907** (2.865)	17.559** (4.802)
Share of Trade in GDP		-3.743** (0.765)	-3.677** (0.540)	-2.256** (0.355)	-2.272* (0.768)	-2.451** (0.407)	-1.594** (0.452)
# of Observations	13	13	13	13	13	13	13
Adjusted $R^2$	0,717	0,957	0,964	0,909	0,873	0,932	0,831

Notes: White heteroskedastic-consistent standard errors are in parentheses. \*\*, \* and # denote significant at the 1%, 5% and 10% level, respectively

As I am particularly interested in the relationship between openness and urban concentration, regression (2) includes the average share of trade in GDP as an additional independent variable. Three points are noteworthy. First, the estimated coefficient on openness is negative and statistically highly significant, indicating that a rise in the share of trade in GDP by 1 percent is associated with a reduction in the size of the

largest city by about 3.7 percent. Holding population constant, this result provides some first support for *Krugman* and *Elizondo's* (1996) hypothesis of a negative relationship between trade openness and urban concentration. Moreover, the estimated effect is considerably larger than *Ades* and *Glaeser's* (1995) finding of -0.6 for a sample of 85 countries.

Second, with openness included, the overall estimation results improve considerably. The fit of the regression is excellent, with an adjusted  $R^2$  of 0.96. This suggests that openness captures a large share of the difference in the size of the largest cities. As there is a well-known association between country size and openness, with larger countries being relatively less open to trade (see, for example, *Frankel* 1997, p. 57), this finding itself is not surprising. The surprising fact is rather that this expected pattern is apparently mostly captured by the openness measure in the sample, while the coefficients on both land area and urbanized population are statistically less significant.

Finally, also the coefficients on most of the other explanatory variables become statistically significant in regression (2). The estimated coefficient on the capital city dummy is now statistically highly significant (at the 1% level), indicating that main cities are on average 45 percent larger if they are also capital cities<sup>10</sup> – a magnitude which is comparable to *Ades* and *Glaeser's* result of an elasticity of 0.42. Thus, even in a sample which includes almost exclusively countries with an established democratic political system (with the exceptions of Portugal and Spain in 1970), I find support for *Ades* and *Glaeser's* claim that political power attracts population. Interestingly, I also replicate their somewhat surprising empirical result that of the two population controls only the log of non-urbanized population is statistically significant. The positive coefficient of about 0.5 indicates that more populous countries have larger central cities, with urban areas growing about half as much as their countries' (non-urban) populations. In contrast to this finding, the coefficient on the log of land area (which is statistically significant at the 5% level) takes the wrong sign, suggesting that geographically larger countries tend to have **smaller** cities. As the partial correlation between country size and main city population, however, is positive (with a coefficient of 0.614), the estimated coefficient on land area possibly reflects a spurious correlation after controlling for the effects that more populous and more closed economies have larger metropolises. As expected, the share of the labor force outside of agriculture is

---

<sup>10</sup> In my sample, three out of 13 cities are not the capital of their respective country (Antwerp/Belgium, Hamburg/Germany, Zurich/Switzerland). This is a considerably larger share than in *Ades* and *Glaeser's* study who report that 77 of the 85 large cities in their sample are also capitals.

positively associated with the size of the main city. A 1 percent increase in the fraction of the population that is not working in the agricultural sector tends to raise the size of the largest city by about 16 percent. At first sight, this elasticity appears to be implausibly large. However, given that the sample mainly consists of highly industrialized countries with only small deviations from an average non-agricultural labor share of 0.95, this result is not unrealistic. Contrary to expectations, the estimated coefficient on per capita income is actually negative, indicating that a 1 percent rise in GDP per capita decreases the size of the main city by about 2.9 percent. It is worth remembering, however, that the share of labor outside of agriculture already controls for a country's state of industrial development, suggesting that this variable captures even more-than-proportionally the expected positive effect of economic development on city size. This interpretation is consistent with *Ades and Glaeser's* (1995) findings who report that in their sample the coefficient on income (although positive) loses size and significance whenever they also control for the non-agricultural labor share. Moreover, the results might be distorted by an outlier. The country with the highest per capita income in the sample, Switzerland, has on average the second smallest main city.

As the most obvious problem in this type of regression (using period averages) is the small sample size, I provide a number of robustness checks. In a first attempt, I reduce the number of explanatory variables to increase the degrees of freedom. In particular, I drop the log of urbanized population (regression 3)<sup>11</sup> which has been found to have no predictive power for the population size in the main city and the log of land area (regression 4) from the list of explanatory variables. As expected, departing from *Ades and Glaeser's* standard set-up does not affect the basic results. In fact, the only slight reduction in the adjusted  $R^2$  from 0.96 to 0.91 suggests that not much information is lost. The most notable change, then, is the fall in magnitude in the estimated coefficient on openness by more than one-third. This apparent sensitivity of the estimated elasticity to the exclusion of land area supports the intuition that the implausible negative coefficient on land area<sup>12</sup> is indeed the result of a misspecification so that I use regression (4) as my preferred specification.

---

11 I also tried running regressions with the country's total population outside the main city as an explanatory variable instead of entering a country's non-urbanized and urbanized population separately. Merging both initial measures of a country's population size into a single variable, however, did not improve the results.

12 *Junius* (1999) also finds a negative relationship between country size and urban concentration and explains this result by arguing that a large land area apparently increases the probability of forming several metropolises.

Columns (5) to (7) present separate regressions for the years 1970, 1980 and 1990 to make sure that the previous findings for period averages are not driven by a specific year. As shown, the basic regression results hold for all individual years. Most notably, the estimated coefficient on the share of trade in GDP is always statistically significant, but varies between -1.6 and -2.5 across different years.

The results in table 3 deal even more explicitly with the issue of small sample size. In particular, I apply different pooling techniques to take fuller account of the available information for the time period from 1970 to 1990. In a first step, regressions (1) and (2) combine observations from the different years, using the method of a seemingly unrelated regression (SUR). Specifically, I estimate a system of three year-specific equations with separate intercepts for each year, but impose a constancy restriction on the remaining parameters. Having increased the number of observations to 39, the basic results are qualitatively unchanged from the comparable OLS regressions with averaged data. The estimated coefficient on the variable of interest, openness, turns out to be negative and statistically highly significant at the 1% level. The magnitude of the estimated elasticity, however, is lower at about -1.4.

Columns (3) and (4) report regressions with pooled data for 1970, 1980 and 1990 in which dummy variables for 1970 and 1980 (not shown) capture time effects. Although the overall results appear to be somewhat weaker, this set-up basically confirms previous findings. As before, the share of trade in GDP enters negatively and statistically highly significant. It is also interesting to note that in both specifications which use a larger number of observations the coefficient on the log of land area is far from significant, supporting my initial decision to ignore this variable in the preferred specification of regressions examining period averages and yearly data. Moreover, the regressions strongly suggest that the previously implausible result of a negative coefficient on GDP per capita is not robust. Contrary to the estimates for averaged data, per capita income is neither in SURs nor in OLS regressions with pooled data statistically different from zero.

In a further exercise in sensitivity analysis, I drop observations for Austria and Germany to make sure that the results are not affected by distorted data due to border changes.<sup>13</sup>

---

<sup>13</sup> Even a very crude analysis suggests that persistence effects might play an important role. While Austria has about the same population size as Belgium, Portugal and Switzerland, its largest city, Vienna, has on average four times as much inhabitants as the main cities in other countries of comparable size.



Columns (5) to (7) of table 3 display the results for different estimation techniques. As shown, the estimated coefficients are virtually unchanged from the regression specifications using the larger sample of countries so that I can rule out that the results are driven by distorted or corrected data.

**Table 3: Alternative Estimation Techniques and Sample Sizes**

Dependent Variable:	Full Sample				Without Austria and Germany		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of Population in Main City	SUR 1970-90	SUR 1970-90	Pooled 1970-90	Pooled 1970-90	Avg. 1970-90	SUR 1970-90	Pooled 1970-90
Capital City Dummy	0.644** (0.157)	0.604** (0.165)	0.581** (0.145)	0.573** (0.145)	0.345 (0.234)	0.766** (0.175)	0.691** (0.204)
Log of Non-Urbanized Population	0.401** (0.094)	0.555** (0.074)	0.414** (0.089)	0.565** (0.051)	0.503** (0.079)	0.520** (0.063)	0.555** (0.046)
Log of Urbanized Population Outside the Main City	0.171 (0.072)		0.149# (0.082)				
Log of Land Area	-0.029 (0.076)		-0.060 (0.082)				
Log of Real GDP per Capita	-0.068 (0.268)	0.039 (0.319)	-0.471 (0.580)	-0.460 (0.525)	-2.024# (0.841)	0.054 (0.264)	-0.290 (0.639)
Share of the Labor Force Outside of Agriculture	1.857* (0.859)	2.663* (1.093)	3.557 (2.996)	5.217* (2.100)	11.656* (3.074)	2.434* (0.919)	4.135 (2.567)
Share of Trade in GDP	-1.297** (0.341)	-1.399** (0.306)	-1.566* (0.737)	-1.540** (0.406)	-2.353* (0.633)	-1.263** (0.273)	-1.280* (0.515)
# of Observations	13 x 3	13 x 3	39	39	11	11 x 3	33
Adjusted R <sup>2</sup>	0.769, 0.828, 0..500	0.822, 0.858, 0..565	0,847	0,838	0,941	0.854, 0.891, 0..516	0,860

Notes: White heteroskedastic-consistent standard errors are in parentheses. \*\*, \* and # denote significant at the 1%, 5% and 10% level, respectively. The seemingly unrelated regression (SUR) includes year-specific intercepts that are not reported here. All other regressions also include a constant which is not reported. The pooled OLS regression additionally include dummy variables for 1970 and 1980

To check the robustness of the finding of a negative association between openness and the size of the largest city, I experiment with various extensions of my benchmark specification. In particular, I follow *Ades and Glaeser (1995)* in entering a number of

additional control variables. Table 4 shows the results. Note that these regressions include the complete set of independent variables considered in regression (4) of table 2, but the table only reports the estimated coefficients on the two variables of particular interest to avoid clutter. Regressions (1) and (2) then examine the trade-city size connection in more detail by including additional measures on the openness of a

**Table 4: Additional Explanatory Variables**

Dependent Variable:	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Log of Population in Main City							
Share of Trade in GDP	-2.086** (0.313)	-2.235** (0.382)	-2.844** (0.388)	-2.788** (0.259)	-2.129** (0.296)	-2.172** (0.317)	-2.329** (0.376)
<i>Other Openness Measures</i>							
Ratio of Import Duties to Imports	4.005# (1.926)						
Average Tariff Rate on Imports		6.262 (6.877)					
<i>Internal Infrastructure</i>							
Log of Railway Density, 1970-90			0.313** (0.100)				
Log of Passenger Cars, 1970-90				-1.822* (0.519)			
<i>Political Variables</i>							
Dictatorship Dummy, 1970					0.381 (0.314)		
Number of Cabinet Changes, 1970						-0.020 (0.025)	
Number of Changes in Effective Executive, 1970							0.012 (0.041)
# of Observations	13	13	13	13	13	13	13
Adjusted R <sup>2</sup>	0,914	0,912	0.957	0,937	0,906	0,900	0,896

Notes: Other regressors not shown in the table are a constant, a capital city dummy, the log of non-urbanized population, the log of real GDP per capita, and the share of the labor force outside of agriculture. White heteroskedastic-consistent standard errors are in parentheses. \*\*, \* and # denote significant at the 1%, 5% and 10% level, respectively.

country. In a first step (column 1), I add the average ratio of import duties to total imports. This variable is intended to capture the impact of protectionist trade policies on the size of the largest city. Specifically, given a negative relationship between openness

and main city size, I would also expect that a larger share of tariff revenues is associated with a more populous central city. The estimated coefficient on the import duties/imports ratio is indeed positive and significant so that there is some support for the intuition. Two points, however, are noteworthy. First, with a coefficient which is significant only at a 10% level of confidence, the variable has only weak predictive power. Second, the estimated coefficient on openness is largely unaffected and remains highly significant. A possible explanation for this result is the generally low level of tariffs in industrialized countries. Moreover, as most of the countries in my sample are members of a common customs union, the European Union, a large share of their international trade is actually free of tariffs. Hence, differences in the ratio of import duties to imports do not necessarily indicate different trade policies, but may mainly reflect differences in the regional and commodity structures of the countries' international trade position.

A better way, then, to discriminate between different trade policies may be to focus on actual tariff rates as a measure of a government's intention to protect the domestic market. The result (column 2), however, is even weaker than the finding for effective protection. Although the average tariff rate also enters positively, the coefficient is insignificant while the coefficient on the share of trade in GDP remains unchanged.

In regressions (3) and (4), I control for the internal infrastructure of a country. It has often been argued that internal transport costs are an important determinant of the national urban structure, although the sign of the relationship is a priori unclear. My results do support this ambiguity. On the one hand, I find a highly significant, positive coefficient on the log of railway density, indicating that countries with a better developed railway infrastructure have larger cities. On the other hand, the log of passenger cars per capita enters negatively and significant, indicating that countries which are better equipped with passenger cars measured in relation to the country's total population tend to have on average smaller main cities. Most notably for my point of interest, however, the estimated coefficient on the share of trade in GDP remains quantitatively unchanged and is statistically highly significant for both transportation variables.

Finally, also controlling for the political structure of a country does not change the finding of a robust negative relationship between openness and city size. Column (5) includes a dictatorship dummy which takes a value of one if the country in question had

a dictatorial regime in 1970. In effect, the dummy assigns a value of one to the two Southern European and least developed countries in my sample, Portugal and Spain. However, the estimated coefficient is statistically not significant, even though it has the correct sign. Consistent with their theoretical discussion, *Ades and Glaeser's* (1995) find a positive coefficient on a dictatorship dummy, indicating that countries with a political system which ignores the political rights of their citizens have on average larger cities.

As my sample mostly comprises countries with a democratic political structure, columns (6) and (7) add variables which reflect the **stability** of the political system. Both controls, however, enter with insignificant coefficients which, in addition, take different signs, suggesting that in my sample political instability has no impact on city size.

In sum, the results largely confirm *Ades and Glaeser's* (1995) findings for a subsample of 13 European countries. In fact, given the immense statistical problems raised by the small size of my sample, the results are surprisingly robust. Most notably, I find a tight and robust negative relationship between the openness of a country and the size of its largest city. Moreover, there is also strong support for the hypothesis that main cities are on average larger if they are also capital cities.

## 4.2 More Years of Data

While in a first set of estimates I basically tried to replicate *Ades and Glaeser's* (1995) empirical results, I now begin to explore the advantages of my data set. In particular, I use the availability of historical data to check whether the negative empirical association between openness and main city size is robust over time. In the previous section, results applying data for separate years have shown that the estimated coefficient on openness consistently turns out to be significant at conventional levels of confidence. Moreover, the coefficient is in all cases clearly negative and economically large in magnitude. Nonetheless, there is some variation in the economic and statistical significance of the coefficient on the share of trade in GDP across individual years.

As I have data going back to 1870, I repeat the regressions for the individual years 1970, 1980, and 1990 for each of the earlier decades using, for comparability, the same set of control variables. Table 5 displays the results. At a first look, there is – as before –

considerable variance in the results for separate years. Starting with the estimates for 1870<sup>14</sup>, all coefficients have the expected, theoretically correct sign and are statistically highly significant. Most notably, openness has a measurable negative impact on the size of the largest city with an estimated elasticity of -7.6 which is more than three times higher than that found for later years. With an adjusted R<sup>2</sup> of 0.99, the fit of the regression is excellent.

**Table 5 More Years of Data**

Dependent Variable: Log of Population in Main City	(1) 1870	(2) 1880	(3) 1890	(4) 1900	(5) 1910	(6) 1920	(7) 1930	(8) 1940	(9) 1950	(10) 1960
Constant	-16.772** (2.391)	-8.326 (4.661)	-9.972 (5.661)	-6.066 (4.838)	-4.165 (2.915)	-2.688 (6.798)	-8.243# (3.891)	-2.360 (5.620)	1.679 (5.434)	1.290 (4.706)
Capital City Dummy	0.657** (0.065)	0.851# (0.421)	0.872* (0.329)	0.882* (0.295)	0.902** (0.241)	1.048** (0.249)	0.968** (0.207)	1.114** (0.318)	1.271* (0.456)	1.019* (0.291)*
Log of Non-Urbanized Population	0.411** (0.052)	0.622* (0.203)	0.553* (0.194)	0.701** (0.172)	0.698** (0.139)	0.850** (0.146)	0.869** (0.151)	0.580* (0.171)	0.569* (0.219)	0.546* (0.188)
Log of Real GDP per Capita	3.143** (0.453)	1.332 (0.773)	1.774 (1.193)	0.869 (0.799)	0.592 (0.606)	0.011 (1.044)	0.867# (0.400)	0.635 (0.577)	0.031 (0.501)	0.154 (0.420)
Share of the Labor Force Outside of Agriculture	3.132* (0.773)	2.725 (1.922)	2.279 (1.388)	1.559 (1.407)	2.137# (1.117)	3.335 (1.991)	0.365 (1.734)	1.583 (1.912)	3.757 (2.501)	3.413 (1.938)
Share of Trade in GDP	-7.600* (0.740)*	-1.793 (1.154)	-2.920 (1.950)	-0.156 (0.263)	-0.101 (0.180)	0.119 (0.718)	-0.236 (0.781)	-0.816 (1.268)	-1.869 (1.489)	-1.709 (1.099)
# of Observations	10	11	12	13	13	13	13	13	13	13
Adjusted R <sup>2</sup>	0,990	0,840	0,876	0,868	0,842	0,717	0,685	0,642	0,597	0,750

Notes: White heteroskedastic-consistent standard errors are in parentheses. \*\*, \* and # denote significant at the 1%, 5% and 10% level, respectively.

14 As I have no data on the structure of the labor force in Portugal in 1870 and 1880 and the non-urbanized population in the Netherlands (1870-1890), Portugal (1870-1880) and Spain (1870), columns (1) to (3) include less than 13 observations.

In the following decades, the overall results are generally weaker with, for instance, the coefficients on per capita income and the non-agricultural share of the labor force losing statistical significance. Also the estimated coefficient on the trade-to-GDP ratio is smaller in size and becomes statistically insignificant. Although the coefficient consistently turns out to be negative, with the exception of 1920, the estimated elasticity is typically well below 0.5 and, thus, considerably smaller than in the estimates for the time period from 1970 to 1990. Beginning in 1880, then, only two controls enter consistently with statistically significant coefficients. The capital city dummy is positive and with values around 1.0 somewhat larger in magnitude than in the regressions for the period from 1970 to 1990. Not surprisingly (given previous results), also the coefficient on the non-urban population of a country turns out to be highly significant, taking the expected positive values.

The results again change after 1940. The most interesting difference is that the coefficient on the variable of interest, openness, although still not statistically significantly different from zero, more than doubles in magnitude. The value of about -1.8 is not very different from the significant estimates for later years.

In summary, the year-specific estimates provide some interesting insights. First, even though differences in results for separate years should be interpreted with care due to the small sample size, there is convincing evidence that the initial regression results for the period from 1970 to 1990 are not robust over time. In particular, I find considerable variation in the predictive power of some controls. As both the magnitudes and the significance levels of those variables, however, are often robust for a number of consecutive years, this variation is apparently not the result of measurement problems or data inconsistencies, but rather seem to reflect structural changes. It is also interesting to note that in its actual specification the fit of the regression consistently decreases in the first half of this century. While the adjusted  $R^2$  is above 0.86 in 1900, the controls explain less than 60% of the variation in the log of main city population in 1950.

Second, the finding of a negative association between trade exposure and the size of a country's largest city is confirmed for 1870, but does not hold for later years. This means that the share of trade in GDP enters only significant in four separate years (of which three – 1970, 1980 and 1990 – make up the time period which is usually examined in other studies!) and, thus, in just one third of my total time series of 120

years of data. This result, however, forcefully questions the empirical robustness of a relationship between openness and main city size.

A potential explanation for the observed differences in the results for individual years are boundary changes. Remember that after World War II, for example, I use data based on the former West German territory while before I have used information covering the former Deutsches Reich territory. Moreover, from this time on, I ignore Berlin as the largest German city and consider instead Hamburg as the main city. Despite those ad hoc conventions, however, I find, for instance, a positive and statistically significant coefficient on the capital city dummy which is robust for the complete time period from 1890 to 1990. In combination with earlier results showing the robustness of estimated coefficients after the exclusion of Austria and Germany, this suggests that the distorting effect of allowing for boundary changes is rather limited.

### 4.3 Other Concentration Measures

In a third extension, I examine whether the negative relationship between openness and main city size is robust for different measures of urban concentration. In particular, I regress three alternative primacy measures on my standard set of explanatory variables.<sup>15</sup> Such a procedure is justified for at least three reasons. First, as I have argued that all controls are potential economic determinants of urban concentration, there is no need to change the regression specification. Accordingly, I would also expect exactly the same relationships. Second, other recent studies focusing, for example, on the share of a country's largest city in urban population such as *Moomaw* and *Shatter* (1996) basically include the same independent variables or slight variants of it. Finally, using an unchanged set of controls has the advantage that the results are directly comparable with my previous findings.

Table 6 shows the results. The estimates vary considerably across the different specifications. In fact, not only the signs and significance levels of single coefficients differ, but also the empirical fit of the regression varies markedly. While the set of controls explains about 70% of the variation in the share of the main city in total population and the *H* concentration measure, the adjusted  $R^2$  for the ratio of the

---

<sup>15</sup> I have again replaced the log of non-urbanized population with the log of a country's total population which slightly increases the fit of the regressions.

population size between a country's two largest cities is only 0.1. This finding supports the hypothesis that different measures of urban primacy indeed cover different aspects of the urban structure.

**Table 6: Other Concentration Measures**

Dependent Variable:	(1) Log of Share of Main City in Total Urban Population	(2) Log of Ratio of Main City to Second Largest City	(3) H Concentration Measure
Constant	22.978* (8.438)	4.036 (10.272)	37.996* (12.948)
Capital City Dummy	0.232 (0.157)	0.145 (0.280)	0.091 (0.352)
Log of Total Population Outside the Main City	-0.720** (0.168)	-0.270 (0.169)	-1.058** (0.240)
Log of Real GDP per Capita	-1.266 (1.194)	0.302 (1.293)	-2.645 (1.700)
Share of the Labor Force Outside of Agriculture	-0.645 (6.154)	-1.542* (5.937)	1.895 (7.936)
Share of Trade in GDP	-1.587 (0.870)	-0.818 (0.974)	-2.155 (1.262)
# of Observations	13	13	13
Adjusted R <sup>2</sup>	0,717	0,102	0,687

Notes: White heteroskedastic -consistent standard errors are in parentheses. \*\*, \* and # denote significant at the 1%, 5% and 10% level, respectively. All data are averages for the period from 1970 to 1990.

In the light of this observation, then, it is somewhat surprising that the findings for the share of trade in GDP are quite unequivocal. In all three specifications, the trade/GDP ratio enters negatively but with an estimated coefficient that is not statistically different from zero. This finding provides further evidence that a negative association between openness and urban concentration is empirically not robust.



#### 4.4 Full Time Period, 1870-1990

In a next step, I turn to the question whether there is a statistically significant association between openness and urban concentration, exploring the full time period for which I have data. This exercise might be of particular interest for at least two reasons. First, reviewing the results of the cross-country regressions, the small size of my sample might indeed be a serious problem. In fact, it is quite striking that I find almost consistently a negative coefficient on the share of trade in GDP, but with varying degrees of statistical significance. Yet this fragility of the empirical results does not necessarily reflect fundamental differences in the relationship between trade openness and urban concentration across individual years. With only 13 observations, it may also be simply due to some mismeasured data in individual years. Extending the time period then and, thereby, increasing the number of observations would possibly minimize the impact of any distorted data.

Second, even if the results for various specifications are not distorted and the different significance levels illustrate 'true' variations in the importance of trade policies for the degree of urban centralization, it would be interesting to know which effect dominates and, thus, whether the finding of a statistically significant negative relationship also holds for the entire time period of 120 years.

The statistical results are in table 7. Each column reports the result of a pooled regression for the period from 1870 to 1990. As before, the parameter estimates vary considerably across different specifications. Most notably, the development controls take on changing signs for different concentration measures. Also the explanatory power of the controls varies markedly across the regressions, ranging from 0.13 to 0.85.

Given this general fragility, then, the estimated coefficients on openness deliver a somewhat more convincing picture as they are consistently negative. But only for two of the four different dependent variables, the trade/GDP ratio enters with a statistically significant (at the 5% level of confidence) coefficient.<sup>16</sup> Thus, the much larger number of observations hardly affects previous findings. If anything, they confirm that the association between trade openness and urban concentration is not robust, depending on the chosen measure of urban concentration.

---

<sup>16</sup> *Nitsch* (2001b) finds that the coefficient loses its significance in a panel estimation with fixed country effects.

**Table 7: Full Time Period, 1870-1990**

Dependent Variable:	(1) Log of Population in Main City	(2) Log of Share of Main City in Total Urban Population	(3) Log of Ratio of Main City to Second Largest City	(4) Log of H Concentration Measure
Constant	-4.869** (1.072)	7.698** (1.266)	-19.799* (8.748)	1.606** (0.278)
Capital City Dummy	0.937** (0.072)	0.693** (0.074)	2.151** (0.297)	0.027** (0.009)
Log of Non-Urbanized Population	0.751** (0.034)			
Log of Total Population Outs. Main City		-0.472** (0.030)	-0.069 (0.179)	-0.036** (0.005)
Log of Real GDP per Capita	0.518** (0.142)	-0.329# (0.172)	3.404** (1.128)	-0.138** (0.035)
Share of Labor Force Outs. Agriculture	2.183** (0.396)	-0.320 (0.464)	-5.593 (3.678)	0.083 (0.078)
Share of Trade in GDP	-0.317* (0.154)	-0.353* (0.164)	-0.779 (0.568)	-0.033 (0.023)
1880	0.228 (0.192)	0.118 (0.178)	-0.195 (1.029)	0.011 (0.023)
1890	0.388* (0.182)	0.280 (0.188)	-0.330 (1.080)	0.053 (0.035)
1900	0.485* (0.187)	0.406* (0.186)	-0.779 (1.049)	0.074* (0.035)
1910	0.544** (0.176)	0.520** (0.190)	-0.853 (1.084)	0.097** (0.037)
1920	0.685** (0.189)	0.592** (0.187)	-0.693 (1.160)	0.116** (0.038)
1930	0.634** (0.206)	0.709** (0.192)	-1.416 (1.146)	0.143** (0.036)
1940	0.453* (0.208)	0.569* (0.222)	-2.172* (1.064)	0.131** (0.035)
1950	0.458* (0.212)	0.654** (0.222)	-2.149# (1.099)	0.146** (0.035)
1960	0.244 (0.219)	0.801** (0.241)	-3.100** (1.133)	0.198** (0.042)
1970	-0.061 (0.247)	0.856** (0.284)	-4.476** (1.298)	0.235** (0.048)
1980	-0.248 (0.264)	0.824** (0.302)	-5.208** (1.394)	0.256** (0.055)
1990	-0.267 (0.291)	0.825** (0.315)	-5.659** (1.527)	0.273** (0.060)
# of Observations	163	163	167	163
Adjusted R <sup>2</sup>	0,849	0,727	0,136	0,383

Notes: White heteroskedastic-consistent standard errors are in parentheses. \*\*, \* and # denote significant at the 1%, 5% and 10% level, respectively.

#### 4.5 Changes in Urban Concentration

Finally, I explore the time series dimension of my sample in more detail. As noted before, there are not only considerable differences in levels across the European economies, but there is also a broad variation across time. Therefore, it might be interesting to check in more detail whether the hypothesis of a negative correlation between an economy's exposure to international trade and the country's degree of population centralization holds for changes across time.

Besides of simply testing the results of the cross-country regressions, however, this specification may also yield a number of additional insights. For one thing, the analysis of changes provides a somewhat different perspective on the relationship between openness and urban concentration. In particular, it allows to examine a related version of the argument of a negative association, namely whether, holding population growth constant, central cities grow less quickly when economies become more open and vice versa.

But there are econometric implications as well. First, previous results may suffer from omitted variables. In the first differenced version, however, problems related to the measurement or omission of variables will not bias the results if these unobserved country characteristics are constant over time. Second, a country's exposure to trade may also be a function of urban concentration: in a highly centralized economy, domestic firms might benefit from a cost advantage over foreign suppliers which would negatively affect the amount of foreign trade. This raises the potential problem of endogeneity. Timing, i.e., the analysis of the correlation between initial variables and later changes, might then provide an (admittedly crude) check of causality.

Table 8 presents the results of simple pooled regressions in the first-difference specification with the estimations running from 1880 to 1990 (the first period being used in data construction). Across all concentration measures, the precision of the coefficients, as measured by the standard errors, is lower than in the traditional level specification. For none of the basic explanatory variables, I find a result which is robust for the four different dependent variables. Either the degrees of statistical significance vary considerably or the coefficients even take on changing signs. Given these generally weak results then, it is not surprising that also the estimates on openness give no convincing picture. The coefficient is in no case statistically different from zero and actually has the wrong sign. The consistently positive estimate suggests that, holding other effects constant, central cities *grew* on average when economies became more open to trade. This finding, however, is perhaps driven by a decrease in urban concentration in the inter-war period when also the volume of trade contracted. Dividing the entire time period into three sub-periods 1880-1910, 1920-1940, 1950-1990 and running separate regressions supports this hypothesis (results not shown). The change in openness enters negatively (and weakly significant) in the first period, then turns positive and statistically highly significant in the inter-war years, and remains positive (but insignificant) in the period from 1950 to 1990.

**Table 8: Changes in Urban Concentration, 1870-1990**

Dependent Variable:	(1)	(2)	(3)	(4)
	$\Delta$ Log of Population in Main City	$\Delta$ Log of Share of Main City in Total Urban Population	$\Delta$ Log of Ratio of Main City to Second Largest City	$\Delta$ Log of H Concentration Measure
Constant	0.219** (0.048)	-0.002 (0.068)	-0.075 (0.049)	0.062 (0.110)
$\Delta$ Capital City Dummy	0.244 (0.173)	0.273* (0.158)	0.112# (0.066)	0.385# (0.220)
$\Delta$ Log of Non- Urbanized Population	0.171** (0.059)			
$\Delta$ Log of Total Population Outs. Main City		-0.615** (0.080)	0.142* (0.061)	-1.303** (0.124)
$\Delta$ Log of Real GDP per Capita	0.241 (0.191)	0.200 (0.162)	-0.057 (0.147)	0.328 (0.228)
$\Delta$ Share of Labor Force Outs. Agriculture	0.531 (0.447)	0.092 (0.345)	0.188 (0.417)	-0.287 (0.546)
$\Delta$ Share of Trade in GDP	0.066 (0.048)	0.058 (0.050)	0.024 (0.070)	0.066 (0.072)
1880-90	0.005 (0.069)	0.041 (0.083)	0.076 (0.088)	0.081 (0.137)
1890-00	-0.056 (0.058)	0.004 (0.072)	0.010 (0.071)	-0.012 (0.118)
1900-10	-0.103# (0.057)	0.009 (0.069)	0.055 (0.053)	0.014 (0.110)
1910-20	-0.056 (0.069)	0.096 (0.084)	0.149* (0.067)	0.112 (0.129)
1920-30	-0.151* (0.071)	-0.009 (0.081)	0.062 (0.061)	-0.051 (0.127)
1930-40	-0.171* (0.067)	-0.042 (0.077)	0.010 (0.078)	-0.108 (0.116)
1940-50	-0.206* (0.091)	-0.079 (0.089)	0.122* (0.058)	-0.172 (0.129)
1950-60	-0.247** (0.082)	-0.100 (0.089)	0.036 (0.059)	-0.116 (0.181)
1960-70	-0.340** (0.086)	-0.221* (0.088)	-0.026 (0.082)	-0.389** (0.134)
1970-80	-0.364** (0.062)	-0.237** (0.075)	-0.011 (0.061)	-0.364** (0.132)
1980-90	-0.224** (0.082)	-0.098 (0.094)	0.151* (0.068)	-0.312* (0.127)
# of Observations	150	150	154	150
Adjusted R <sup>2</sup>	0,297	0,426	0,020	0,510

Notes: White heteroskedastic-consistent standard errors are in parentheses. \*\*, \* and # denote significant at the 1%, 5% and 10% level, respectively.

To provide a rough test for causality, I slightly change the simple first-difference specification and replace the change in openness by the share of trade in GDP at the beginning of the decade as explanatory variable. I now examine whether openness is a statistically significant predictor of the change in urban concentration in the subsequent decade. The results, displayed in table 9, are not encouraging. For each of the concentration measures, the coefficient on external trade takes the correct (negative) sign but is not statistically significant at conventional levels of confidence. In regressions not reported here, I have also experimented with longer (20-year) intervals.

The basic results were virtually unchanged. The negative coefficient on the trade-to-GDP ratio slightly increases in magnitude, but remains statistically insignificant.

**Table 9: Openness and Changes in Urban Concentration, 1870-1990**

Dependent Variable:	(1) Δ Log of Population in Main City	(2) Δ Log of Share of Main City in Total Urban Population	(3) Δ Log of Ratio of Main City to Second Largest City	(4) Δ Log of H Concentration Measure
Constant	0.234** (0.050)	0.009 (0.070)	-0.064 (0.051)	0.072 (0.109)
Δ Capital City Dummy	0.240 (0.174)	0.269# (0.158)	0.110# (0.066)	0.381# (0.221)
Δ Log of Non- Urbanized Population	0.169** (0.059)			
Δ Log of Total Population Outs. Main City		-0.620** (0.080)	0.144* (0.060)	-1.310** (0.125)
Δ Log of Real GDP per Capita	0.239 (0.195)	0.201 (0.165)	-0.063 (0.147)	0.330 (0.232)
Δ Share of Labor Force Outs. Agriculture	0.508 (0.433)	0.070 (0.335)	0.176 (0.407)	-0.312 (0.535)
Share of Trade in GDP	-0.021 (0.049)	-0.011 (0.053)	-0.022 (0.048)	-0.007 (0.064)
1880-90	-0.001 (0.070)	0.035 (0.083)	0.076 (0.090)	0.074 (0.137)
1890-00	-0.060 (0.059)	0.000 (0.073)	0.012 (0.071)	-0.017 (0.119)
1900-10	-0.098# (0.058)	0.012 (0.070)	0.059 (0.052)	0.017 (0.110)
1910-20	-0.064 (0.068)	0.086 (0.085)	0.149* (0.067)	0.099 (0.130)
1920-30	-0.161* (0.072)	-0.020 (0.082)	0.062 (0.058)	-0.064 (0.129)
1930-40	-0.185** (0.067)	-0.055 (0.076)	0.005 (0.076)	-0.123 (0.115)
1940-50	-0.209* (0.091)	-0.081 (0.089)	0.120* (0.058)	-0.174 (0.129)
1950-60	-0.248** (0.083)	-0.100 (0.090)	0.037 (0.060)	-0.117 (0.181)
1960-70	-0.340** (0.089)	-0.222* (0.089)	-0.024 (0.083)	-0.391** (0.136)
1970-80	-0.360** (0.065)	-0.234** (0.076)	-0.007 (0.061)	-0.362** (0.132)
1980-90	-0.228** (0.077)	-0.103 (0.089)	0.153* (0.063)	-0.319* (0.127)
# of Observations	150	150	154	150
Adjusted R <sup>2</sup>	0,295	0,424	0,020	0,509

Notes: White heteroskedastic-consistent standard errors are in parentheses. \*\*, \* and # denote significant at the 1%, 5% and 10% level, respectively.

In sum, there is no evidence for the hypothesis that central cities grow faster in countries with more protectionist trade policies. Once again, this casts doubt on the theoretical

presumption of a robust negative association between openness and urban concentration.

## 5 CONCLUSION

This paper examines the empirical relationship between an economy's exposure to international trade and the degree to which a country's population is concentrated in the largest city. In contrast to previous work (*Rosen and Resnick* 1980, *Ades and Glaeser* 1995, *Moomaw and Shatter* 1996) which is largely based on wide cross-country samples, the analysis in this paper concentrates on only one region of the world, Europe. This focus has the advantage that, on the one hand, not much information is lost as there are still considerable cross-country differences, while, on the other hand, a wide variety of additional information, most notably historical data, is available. In short, I explore data from 13 European countries, covering the time period from 1870 to 1990.

The results can be summarized as follows. First, I am able to replicate *Ades and Glaeser's* finding of a statistically significant negative relationship between openness and the size of the largest city for a European sub-sample for the time period from 1970 to 1990. This finding also proves to be robust to the inclusion of a wide range of controls.

Second, the results are much weaker for other time periods. Examining the entire period from 1870 to 1990 in year-specific regressions, I find a statistically significant negative coefficient on the share of trade in GDP in only four cases and, thus, in less than one-third of the years for which I have data. As this fragility, however, may at least in some part be due to the small sample size, I also implement pooled estimation techniques. These more advanced specifications provide no further insights. In a pooled regression for the whole 120-year period, the estimated coefficient on openness is negative, but the level of statistical significance drops to the 5% level.

Third, the linkage between trade policy and urban centralization is also not robust for alternative measures of urban concentration. Replacing absolute central city size by other concentration measures which are frequently used in the literature yields negative, but insignificant coefficients on the trade-to-GDP ratio.

Fourth, in a more dynamic interpretation of the hypothesis that there is a negative relationship between trade orientation and urban concentration, there is also no evidence that economies become less centralized when a country turns to more open trade policies and thereby increases its exposure to international trade.

In conclusion, the apparent fragility of the empirical association between openness and urban concentration obviously calls into question the validity of *Krugman and Elizondo's* (1996) urban structure model.

## References:

- Ades, Alberto F. and Glaeser Edward L. (1995)*  
Trade and Circuses: Explaining Urban Giants, *Quarterly Journal of Economics* 110 (February): 195-227
- Bairoch, Paul (1988)*  
Cities and Economic Development. Chicago: University of Chicago Press
- Carroll, Glenn R. (1982)*  
National City-Size Distributions: What Do We Know After 67 Years of Research? *Progress in Human Geography* 6 (March): 1-43
- De Long, J. Bradford and Shleife Andrei (1993)*  
Princes and Merchants: European City Growth Before the Industrial Revolution, *Journal of Law and Economics* 36 (October): 671-702
- Frankel, Jeffrey (1997)*  
Regional Trading Blocs in the World Trading System. Washington: Institute for International Economics
- Junius, Karsten (1999)*  
Primacy and Economic Development: Bell Shaped or Parallel Growth of Cities, *Journal of Economic Development* 24 (June): 1-30
- Krugman, Paul R. (1991)*  
Increasing Returns and Economic Geography, *Journal of Political Economy* 99 (June): 483-499
- Krugman, Paul R. (1996)*  
Urban Concentration: The Role of Increasing Returns and Transport Costs, *International Regional Science Review* 19: 5-30
- Krugman, Paul R. and Elizondo Raul Livas (1996)*  
Trade Policy and Third World Metropolis, *Journal of Development Economics* 49 (April): 137-150
- Maddison, Angus (1995)*  
Monitoring the World Economy 1820-1992. Paris: OECD Development Centre
- Marshall, Alfred (1890 [1920])*  
Principles of Economics. 8th Edition. London: Macmillan
- Mitchell, Brian R. (1992)*  
International Historical Statistics Europe 1750-1988. Third Edition. New York: Stockton Press
- Moomaw, Ronald L. and Shatter Ali M. (1996)*  
Urbanization and Economic Development: A Bias toward Large Cities? *Journal of Urban Economics* 40 (July): 13-37
- Nitsch, Volker (2001a)*  
City Growth in Europe. Berlin: Duncker & Humblot (forthcoming).
- Nitsch, Volker (2001b)*  
Does Openness Reduce Urban Concentration? Evidence from 120 Years of European Data, Bankgesellschaft Berlin



*Rosen, Kenneth T. and Resnick Mitchel (1980)*

The Size Distribution of Cities: An Examination of the Pareto Law and Primacy,  
Journal of Urban Economics 8 (September): 165-186

*Wheaton, William C. and Shishido Hisanobu (1981)*

Urban Concentration, Agglomeration Economies, and the Level of Economic  
Development, Economic Development and Cultural Change 30 (October): 17-30

## Appendix

### **Data Sources**

#### City Population

Data on city population are compiled from national statistical yearbooks. In some cases, the data refer to the nearest census date available.

#### GDP

Data on GDP (measured in million 1990 Geary-Khamis Dollars) are taken from *Maddison Angus Monitoring the World Economy 1820-1992* (Paris: OECD) Table C-16. In case of missing data, figures are derived by interpolation (Portugal 1880, 1910, 1920; Spain 1880; Switzerland 1880, 1890). Data for Portugal in 1930 and 1940 are from 1929 and 1938, respectively. Data for Austria before 1920 and for Germany before 1950 are calculated by multiplying GDP per capita taken from *Maddison Angus Monitoring the World Economy 1820-1992* (Paris: OECD) Table D-1 with population taken from *Mitchell Brian R. International Historical Statistics Europe 1750-1988* (New York: Stockton Press) Table A1.

#### Population

Data on population are taken from *Maddison Angus Monitoring the World Economy 1820-1992* (Paris: OECD) Table A-3. Missing data for Spain (1880) and Portugal (1880) are derived by interpolation. Data for Austria before 1920 and for Germany before 1950 are taken from *Mitchell Brian R. International Historical Statistics Europe 1750-1988* (New York: Stockton Press) Table A1.

#### Area

Area data is taken from *Banks Arthur S. Cross-Polity Time-Series Data* (Cambridge: MIT Press) Segment 1 and is converted to square kilometers (by multiplying the original data in square miles with 2.59). The data generally refer to the current territory, except for Austria and Germany.

#### Trade

Data on the share of trade in GDP are calculated from *Mitchell Brian R. International Historical Statistics Europe 1750-1988* (New York: Stockton Press) Tables E1 and J1. In some cases, the data refer to the nearest census date available. Missing values are calculated from *Maddison Angus Monitoring the World Economy 1820-1992* (Paris: OECD) and *Banks Arthur S. Cross-Polity Time-Series Data* (Cambridge: MIT Press) Segment 5. Historical data for Portugal (1870-1940) is derived from *Nunes Ana Bela, Mata Eugenia and Nuno Valerio "Portuguese Economic Growth 1833-1985" Journal of European Economic History 18 (Fall 1989): 291-330.*

### Labor Force Outside of Agriculture

Data on the share of the labor force outside of agriculture, forestry and fishing are calculated from *Mitchell Brian R. International Historical Statistics Europe 1750-1988* (New York: Stockton Press) Table B1. In some cases, the data refer to the nearest census date available. Missing data are derived by interpolation (Belgium 1940, Finland 1890, Germany 1900, Italy 1890, Netherlands 1880, 1940, Norway 1880, Portugal 1920, Spain 1930, Switzerland 1880). Data for Germany 1870 is taken from *Maddison Angus Monitoring the World Economy 1820-1992* (Paris:OECD) Table 2-5 and data for Finland 1870, Norway 1870 and Switzerland 1870 are estimated. Data for 1990 are taken directly from *Mitchell's Brian R.* original source for data from 1968 to 1981, the International Labour Organization's *Yearbook of Labour Statistics* (Geneva: ILO).

### Urbanized Population

Data on the share of urban population are available for the time period from 1950-1990 from the United Nations' *Demographic Indicators of Countries* (New York: UN). Historical data compatible with UN estimates are derived from *Flora Peter* (with *Kraus Franz* and *Pfenning Winfried*) *States, Economy, and Society in Western Europe 1815-1975* (Frankfurt: Campus) volume 2, chapter 3 (except for Portugal and Spain) and from *Pereira Pedro* and *Mata Maria* (eds.) *Urban Dominance and Labour Market Differentiation of a European Capital City: Lisbon 1890-1990* (London: Kluwer) for Portugal. Specifically, comparable data is calculated by using the reported share of population in localities with less than 5,000 inhabitants (for Denmark and Norway less than 2,000; for Italy less than 10,000). In the time period with overlapping data (usually 1950-70), the  $R^2$  between both measures is higher than 0.99 for Austria, Finland, France, Germany, Italy, Norway, Sweden and Portugal, about 0.9 for Belgium and Switzerland, and only for the Netherlands at 0.35. For Spain, I use the share of population in province capitals in total population as a proxy for urbanization. From 1950-70, there is a very strong correlation between this measure and the urbanization rate reported by the UN (with an  $R^2$  of 0.99).

Given the share of urban population, I can then easily calculate the figures for urbanized population outside the main city and for nonurbanized population.

### Ratio of Import Duties to Imports

The ratio is calculated by dividing import duties compiled from various issues of the International Monetary Fund's *Government Finance Statistics Yearbook* (Washington: IMF) Table A line 6.1 by total imports taken from the IMF's *International Finance Statistics Yearbook* (Washington: IMF) line 71.

### Tariff Rate

Data on the tariff rate is taken from *Lee Jong-Wha* "International Trade, Distortions, and Long-Run Economic Growth," IMF Working Paper 92/90, who provides the actual average tariff rate on imported inputs, intermediate, and capital goods in or around 1980.

### Passenger Cars

Data on the usage of passenger cars are calculated by dividing the number of motor vehicles (private cars) in use taken from *Mitchell Brian R. International Historical Statistics Europe 1750-1988* (New York: Stockton Press) Table F6 by total population. Due to missing car data, the number of cars refers in some cases to the nearest census date available.

### Railway Density

Data on railway density are calculated by dividing the length of railway lines open (in kilometers) taken from *Mitchell Brian R. International Historical Statistics Europe 1750-1988* (New York: Stockton Press) Table F1 by land area. Missing data for Austria 1940 is derived by interpolation.

### Dictatorship

Data on the presence of dictatorships are based on *Nicholson Frances (ed.) Political and Economic Encyclopaedia of Western Europe* (Harlow: Longman).

### Cabinet Changes

Cabinet changes refers to the number of times in the preceding decade that a new premier is named and/or 50% of the cabinet posts are occupied by new ministers. The series is compiled from yearly data taken from *Banks Arthur S. Cross-Polity Time-Series Data* (Cambridge: MIT Press) Segment 1. Missing values are replaced by the decade average calculated from the remaining years for which data is available.

### Changes in Effective Executive

Changes in effective executive refers to the number of times in the preceding decade that effective control of the executive power changes hands. The series is compiled from yearly data taken from *Banks Arthur S. Cross-Polity Time-Series Data* (Cambridge: MIT Press) Segment 1. Missing values are replaced by the decade average calculated from the remaining years for which data is available.