September 10, 2003

Trade Openness and Urban Concentration: New Evidence*

Volker Nitsch
Bankgesellschaft Berlin

Abstract

In this paper, I reexamine the empirical relationship between trade openness and urban concentration. In particular, I critically review the available cross-country evidence. I show that previous findings of a negative association between trade openness and the size of a country's largest city are not robust and affected by endogeneity bias. Also I find no evidence that trade liberalization significantly reduces urban concentration.

* This paper draws on ideas initially developed in chapter 4 of my dissertation "City Growth in Europe". I thank my thesis supervisors Dalia Marin and Michael Burda for their encouragement, advice and support. I am also grateful to Luisito Bertinelli, Masahisa Fujita, Henry Overman, Diego Puga, Stephen Redding, and participants at the CEPR conference "The Economics of the Cities" and the ZEW Summer Workshop "Innovation, Location and European Integration" for helpful comments and suggestions. Opinions expressed in this paper do not necessarily represent the views of Bankgesellschaft Berlin.

JEL Code: F15, R12

Keywords: trade liberalization, urban primacy

Address:
Volker Nitsch
Bankgesellschaft Berlin
Volkswirtschaft
Hardenbergstrasse 32
10623 Berlin
Germany
Tel.: (030) 3109-4462
Fax: (030) 3109-5031
E-mail: volker.nitsch@bankgesellschaft.de

I. Introduction

This short paper deals with the question „Is there a relationship between a country's external trade and its internal geography?" or, more precisely, „Does openness matter for urban concentration?" This issue is of interest for at least two reasons.

First, policy-makers and academics are increasingly concerned about excessive concentration; especially in developing countries a disproportionately large share of a country's urban population appears to be concentrated in one or two major metropolitan areas (mega-cities) that strongly dominate the national urban structure. The World Bank (2003, p. 139), for instance, notes in its World Development Report 2003: „[T]he spatial distribution of economic activity in general, and of urban centers in particular, is important to sustainable development. ... Excessive primacy can have real economic efficiency costs to countries." Vernon Henderson (2003) estimates that a deviation from the best primacy level by about 0.1 is associated with a loss in productivity growth by about 0.6% a year. If there is an association between openness to international trade and urban concentration, however, changes in trade policy may be a useful strategy to decentralization.

Second, while for a long time the prevailing view was that the empirical relationship between openness and urban concentration is ambiguous and perhaps positive (with large merchant cities benefiting from a liberal trade regime), there now appears to be a new consensus that the effect is exactly the opposite: *negative* and significant. Paul Krugman (1996, p. 13), for instance, claims that it is one of four stylized facts about urban size distributions that „[m]ore open economies, as measured by the share of exports in gross domestic product, tend to have smaller biggest cities". There are basically three pieces of evidence that support this claim. Several studies by Gordon Hanson document for the case of Mexico that trade liberalization is accompanied by decentralization; the removal of trade barriers initiated in the mid-1980s appears to have contributed to a relocation of the Mexican industry away from Mexico City toward the northern border of the country (see, for instance, Hanson [1998] and Krugman and Hanson [1993]). Krugman and Raul Livas Elizondo (1996) formalize this story; they develop a model in which access to external markets weakens the agglomeration forces inside the economy, making it more likely that the country's internal structure is spatially dispersed. Finally, Alberto Ades and Edward Glaeser (1995) find in a cross section sample of 85 countries that the share of trade in GDP is negatively related to the size of the largest city, holding other things constant.

In this paper, I reexamine the empirical relationship between trade openness and urban concentration. I do so, first, by critically reviewing the cross-country findings by Ades and Glaeser. In particular, I use Ades and Glaeser's approach as a starting point and then modify and extend their empirical framework. I find that the result of a negative relationship between openness and urban concentration is not robust; the estimates of the effect of trade on primacy seem to be sensitive to sample size and the regression specification. In a next step, I focus on trade policy (instead of openness) and explore the effects of changes in a country's trade regime on urban structure. Again, I find no evidence that urban concentration is related to external openness; trade liberalization appears to have no measurable effect on urban primacy. Finally, I take into account the likely endogeneity of trade. The results strongly suggest that ordinary least-squares estimates overstate the effects of trade. If geographic characteristics are used as an instrument for trade, the statistical association between international trade and urban concentration becomes zero.

The remainder of the paper is organized as follows. In the next section, I review Ades and Glaeser's approach and their results. Sections 3 to 6 present the main extensions to their framework, followed by an analysis of the effects of trade liberalization on urban concentration and the instrumental variable approach. Section 9 concludes.

II. Retrospective

The paper by Ades and Glaeser (1995, henceforth AG) is extremely comprehensive. In order to determine the factors behind the centralization of a nation's urban population in a major city, they (i) propose a simple theoretical model (to capture the effects of government and politics on urban primacy), (ii) analyze historical case studies, and (iii) explore cross-country evidence. Another interesting empirical contribution is, for instance, that AG document the impact of political forces on the national urban structure. Here, however, I focus exclusively on their empirical finding of a negative association between trade openness and urban concentration.

AG's empirical strategy is highly intuitive. In particular, they estimate an equation of the form:

$$(1) \qquad \ln(CITY) = \alpha + \Sigma_i \beta_i x_i + \Sigma_j \gamma_j y_j + \Sigma_k \delta_k z_k$$

where the x_i's are scale variables for the population size of the largest city, the (log of) nonurbanized population and the (log of) population in other urbanized areas; the y_j's are the main variables of interest: measures of political stability and trade openness; and z_k is a vector of other controls that have the potential to affect the size of the country's largest city. In their basic specification, the controls include the log of land area, the log of real GDP per capita, and the share of labor force outside of agriculture.

AG estimate this equation using observations spanning four different years: 1970, 1975, 1980 and 1985; the data are averaged (instead of applying panel techniques) to abstract from the question of how lagged values of country characteristics might change current urban concentration.

Concerning the main variable of interest, the results seem to be convincing. The estimated coefficient on the share of trade in GDP is consistently negative. Also, in the benchmark specification, the coefficient is statistically highly significant, with a t-statistic of 2.7. However, AG are also aware that the empirical association between openness and urban concentration is not particularly robust. Of the 7 regressions reported in the paper, the coefficient is insignificant in three perturbations. For instance, the magnitude of the estimated coefficient falls sizably if a Latin America dummy is included. Further, additional tests cast doubt on the causality in these correlations. AG conclude (p. 224): „Trade and cities are connected, but it may be that urban concentration is causing low levels of trade, not that low levels of trade induce concentration."

III. Minor Issues: Methodology and Data

In a first exercise, I seek to replicate AG's benchmark results. The first column of table 1 shows (for comparison) the coefficient estimates of AG's basic specification (taken from their table 4). In column 2, I report the estimates of a comparable regression specification. In particular, I tried to obtain data from the same data sources as AG to stay as closely as possible with their analysis; later I will use data from more actual sources which leads to changes in the country sample since some countries have experienced territorial changes (e.g., Ethiopia, Germany, Yemen). City population data are taken from the United Nation's Prospects of World Urbanization, 1988. The UN population database[1] also provides information on a country's total, urbanized and nonurbanized population. Data on real GDP

[1] Available at http://esa.un.org/unpp.

per capita and the share of trade in GDP are from the Penn World Table mark 5.6. Finally, data on the land area and the share of labor force outside agriculture are obtained from the World Bank's World Development Indicators.

As shown, most of the point estimates are basically identical with AG's results.[2] The size of a country's largest city rises with the country's total population, although only the coefficient on the log of nonurbanized population is economically and statistically significant. Also country size matters; the coefficient estimate implies that an increase in country size by 10 percent increases the population in the main city by about 1.7 percent. Finally, more developed economies tend to have larger central cities. This effect is completely captured by the share of labor force outside agriculture while the coefficient on real GDP per capita is not significantly different from zero.

Turning to the main variable of interest, the coefficient on the share of trade in GDP is negative and almost of the same magnitude as in AG. The elasticity of –0.55 suggests that an increase in the openness ratio by 10 percent reduces the size of the largest city by about 5.5 percent. However, with a t-statistic of 1.7, the coefficient is only marginally significant. While I do not attempt to interpret this result too literally, the weak significance level may be already a first indication of the low robustness of the association between trade openness and urban concentration.

In column 3, I extend the sample of countries. AG restrict their analysis to 85 countries (and even smaller samples) since they later include additional controls. If I enclose all countries for which I have data, my sample covers 115 countries.[3] As shown, increasing the sample size generally improves the precision of the estimates. The coefficient on the capital city dummy becomes statistically highly significant, and the coefficient on the log of real GDP per capita increases in magnitude and significance. The model also explains a slightly higher proportion (86%) of the variation in the cross-sectional data. Most notably for my purposes, however, the share of trade in GDP has now a significantly large effect on the size of the main city; the (absolute) t-statistic rises to 2.7.

In the remaining two columns in table 1, I present analogous estimates for the period 1985 to 2000. City population data are now taken from the World Urbanization Prospects:

[2] My sample covers only 84 countries. Ades and Glaeser note that their sample consists of 85 countries, but their data appendix lists only 84 countries.

[3] It should also be noted that all results concerning the relationship between openness and concentration crucially depend on the exclusion of Hong Kong and Singapore. If the two highly open city states are included, the coefficient on the share of trade in GDP becomes positive.

The 2001 Revision. I also use the Penn World Table mark 6.0 and fill in missing data for the openness ratio with information from the World Bank's World Development Indicators. This update does not affect any of the main results. Most of the coefficient estimates are unchanged from the previous 15-year period. Again, the openness variable enters the regression specification negatively and is economically large and statistically highly significant.

To summarize, the evidence from averaged cross-country data basically supports AG's finding of a negative relationship between trade openness and the size of the largest city. It seems that trade liberalization reduces the average size of the central city. In the following, I will examine the robustness of this result.

IV. Does a City's Geographic Location Matter?

A first extension deals with an issue that is not explicitly addressed in AG but has recently attracted considerable interest, the potential importance of the main city's geographic location within a country. The idea is that an increase in a country's openness to international trade particularly benefits locations close to the border. While these areas have a locational disadvantage in a closed economy, being geographically remote and lacking some rural hinterland, they become increasingly attractive as trade barriers come down, providing good access to international markets. As a result, trade liberalization may increase (rather than decrease) urban primacy if a country's largest city is located on the periphery and thus can be expected to grow through increased trade and commerce. Based on these considerations, Henderson (1996, p. 33) concludes that "the impact of trade on national space is situation-specific, depending on the precise geography of the country."

To control for city location, I experiment with two variables. A first dummy variable takes the value of one if the primate city is a port; this variable is common in the literature (e.g., Henderson [2000]; see also Jordan Rappaport and Jeffrey Sachs [2003]). Port cities, located on the coastal periphery of a country, typically benefit from international trade and thus are often disproportionately large. A second dummy variable focuses more directly on a country's internal geography. It is defined to take the value of one if the main city is located close to the geographic center of the country. Specifically, the variable is constructed by calculating the distance of the city from the approximate geographic center of the country, as given by the latitude and longitude figures in the CIA's World Factbook 2002[4], relative to

[4] Available at http://www.cia.gov/cia/publications/factbook.

country size. If the distance is less than one-half of the square root of the land area of the country, a city is assumed to have a central location.[5] An appendix lists the classification of cities.

The results are reported in table 2. As shown in the first column, there is indeed evidence that primate cities are disproportionately large if they are located on the coast rather than on an interior site. The estimated coefficient on the port city dummy is positive and economically reasonable, though only of moderate statistical significance. Further, similar to Henderson (2000), the raw effect of port is halved if port is interacted with openness. While both coefficients are positive, they are not statistically significant at conventional levels. More generally, however, these extensions leave the estimated openness coefficient basically unaffected with slightly higher standard errors. For the central location dummy, the results are even less satisfactory. The coefficients are insignificant and even change sign for different periods.

Taken together, the empirical results suggest that the geographical location of cities has only a limited impact on the relationship between trade openness and urban concentration. Interaction terms between city location and trade status enter the regression insignificantly, while the standard trade openness coefficient retains its size and only marginally loses statistical significance.

V. Absolute Size vs. Urban Primacy

AG's empirical strategy differs from previous attempts to identify determinants of urban concentration in using the absolute size of the country's largest city as dependent variable. Conceptually, this is not necessarily a problem since the log of urban population outside the main city enters the regression as explanatory variable. Specifically, the estimation equation:

(1a) $\ln(CITY) = \alpha + \beta_1 \ln(URBPOP) + ...$

[5] While this definition is arbitrary, it turns out to work well in practice. The sample of countries is almost evenly divided. According to the measure of relative distance, the five most centrally located primate cities are San José (Costa Rica), Kigali (Rwanda), Windhoek (Namibia), Baghdad (Iraq), and Riyadh (Saudi Arabia); the five most peripheral located cities are Papeete (French Polynesia), Suva (Fiji), Kuala Lumpur (Malaysia), Kuwait City (Kuwait), and Mogadishu (Somalia).

is mathematically equivalent to

(1b) $\ln(CITY/URBPOP^{\beta 1}) = \alpha + ...$,

close to a regression specification that uses urban primacy, the share of the largest city in urban population [i.e., ln(CITY/URBPOP)], as dependent variable.

In the results in tables 1 and 2, however, the point estimates on the log of urban population outside the main city are not different from zero at conventional levels of statistical significance. Therefore, it might be useful to modify the regression specification, using explicitly urban primacy as regressand.

In a first set of regressions reported in table 3, I use the (log of the) share of the main city in urban population outside the main city as dependent variable (i.e., I set $\beta_1=1$). This modification indeed changes the results. Not surprisingly, the coefficient on the log of nonurbanized population becomes negative; in more populous countries a smaller share of the population tends to be concentrated in a central city. Further, the level of economic development (as measured by the share of the labor force outside agriculture) now appears to be uncorrelated with a country's urban concentration. Even more noteworthy is, however, that the coefficient on the variable of interest, the share of trade in GDP, is not significantly different from zero. Trade openness has obviously no measurable effect on urban primacy.

A second set of estimates applies a more conventional regression specification, using the (log of the) share of the main city in *total* urban population as dependent variable.[6] For this limited dependent variable, the coefficient on the trade-to-GDP ratio becomes statistically significant again at conventional levels, but, with t-statistics between 1.9 and 2.5 (for the full country sample), the coefficient is much less precisely estimated than before.

In general, the results highlight the sensitivity of the empirical relationship between openness and urban concentration to the regression specification. A potential explanation for the discrepancy in the results is that primate cities in closed economies may be particularly large relative to cities in open economies, but are probably not that dominant relative to the rest of the national urban system. This hypothesis I will explore next.

VI. Moving Down the City Size Distribution

[6] See, for instance, Henderson (2002).

In this modification, I again use (the log of) absolute city size as dependent variable, but gradually extend the number of cities below the country's largest city. Interaction variables then capture the extent to which the main city is different from the rest of the city size distribution.[7] The main source of data is again the UN's World Urbanization Prospects which compiles information on all cities with more than 750,000 inhabitants, filled in with data from Vernon Henderson's world cities database.[8]

The results are tabulated in table 4. At least three observations are noteworthy. First, extending the sample of cities below the main city appears to reduce the significance of the openness variable. For the period 1970-85, the openness variable even becomes insignificant when more than the countries' largest city are included. Second, although the linkage between trade openness and city size appears to be weaker for cities below the first rank, there is no measurable difference in the link for primate cities. In all perturbations, the interaction term on trade openness and the main city is statistically indistinguishable from zero. This result confirms the findings in the previous section: primate cities in closed economies do not have a particularly dominant position in the national urban system. Finally, the capital city effect which suggests that cities with political functions tend to be disproportionately large is not dependent on a country's central city status.

VII. Liberalization Effects

Having experimented with several variations of the dependent variable, I now modify the external openness measure. Instead of simply defining openness as the share of trade in GDP, I now focus more explicitly on a country's trade policies. This approach has several advantages. For one thing, trade policy (unlike the trade-to-GDP ratio) appears to be completely unrelated to other country characteristics. More importantly, however, this approach deals directly with the policy question of interest, namely: does trade liberalization reduce urban concentration?

To accurately measure a country's overall trade policy stance is a difficult task. Types of trade restrictions vary considerably, ranging from tariff and nontariff barriers to exchange

[7] Nitsch (2001) proposes a similar approach for historical European data.
[8] The data has been gratefully made available by Vernon Henderson at http://econ.pstc.brown.edu/faculty/henderson/worldcities.html.

rate distortions and state monopolies, so that different indicators often give different results.[9] Fortunately, a summary measure is readily available: a dummy variable that classifies countries as open or closed to international trade, constructed by Jeffrey Sachs and Andrew Warner (1995). According to this measure, a country is classified as closed if it displays at least one of the following five characteristics: an average tariff rate of 40% and more; nontariff barriers covering 40% or more of trade; a black market exchange rate that is depreciated by 20% or more relative to the official exchange rate; a state monopoly on major exports; and a socialist economic system. While this methodology is not without criticisms (see Dani Rodrik and Francisco Rodriguez [2000]), Romain Wacziarg and Karen Welch (2002) argue that the dates of trade liberalization derived from both quantitative data and a detailed review of country-specific case studies of reform are a reliable indicator; I use Wacziarg and Welch's corrected and updated data.

In the actual implementation, I run a regression of the form:

$$(2) \qquad \ln(CITY_{it}/URBPOP_{it}) = \gamma_i + \delta\, LIBERAL_{it} + \varepsilon_{it} \ ,$$

where *LIBERAL*=1 if *t* is greater than the year of trade liberalization (and 0 otherwise), and δ is the variable of interest to me. With the inclusion of country fixed-effects, estimates of δ then indicate the within-country variation in urban primacy resulting from a discrete change in trade policy openness.

The results are presented in table 5. I begin with regression results for *LIBERAL* set to 1 when a period of uninterrupted openness began and no reversal of the trade policy reforms occurred, reported in the first part of table 5. As shown, I experiment with several specifications (including a time trend and year dummy variables). In all specifications, δ is indeed negative, indicating that countries that liberalized their trade regime experienced a decline in urban primacy. However, the estimated within-country effect is economically small, averaging about 1%, and rarely statistically significant. In table 5b, *LIBERAL* is defined to also allow for periods of temporary trade liberalization. Not surprisingly, the estimation results are even weaker. While all estimates of δ remain negative, none of the estimates is different from zero at conventional levels of statistical significance. The large variation in the

[9] Reviewing the literature, Andrew Rose (2003) has recently compiled 64 different trade policy indicators.

evolution of urban primacy, unrelated to trade regime, is illustrated in figures 1 and 2 which plot some (carefully chosen) case studies.

VIII. Taking Endogeneity Seriously

A potential problem of the cross-country regression of urban concentration on trade is endogeneity: a country's trade share may be affected by the degree of urban concentration, with large central cities possibly implying less external trade. Also, replacing trade by measures of countries' trade policies may not effectively solve the problem: changes in a country's economic policies may affect urban concentration and trade openness simultaneously and thereby lead to spurious correlation.

A possible solution is the instrumental variable (IV) approach proposed by Jeffrey Frankel and David Romer (1999). In an attempt to identify the effect of trade on income, Frankel and Romer instrument for the actual trade share by constructing a measure of the geographic component of countries' trade. In particular, they estimate a modified gravity equation (where bilateral trade flows are regressed on countries' geographic characteristics) and then aggregate the fitted trade values. This geography-based trade share appears to be a valid instrument. The measure depends only on geographic characteristics. Moreover, the correlation between the actual and the constructed trade share is a reassuringly high 0.62.

Table 6 then reports IV estimates that use Frankel and Romer's (1999) constructed trade share as instrument for trade openness. As shown, treating the trade-to-GDP ratio as endogenous changes the results on the variable of interest considerably. Not only that the estimated coefficient on openness comes down sizably, compared with previous OLS estimates; for all three regression specifications, the effect of openness on primacy is effectively zero. These results provide strong evidence that OLS estimates considerably overstate the effect of trade on urban concentration; there seems to be no causal link between openness and concentration.

IX. Conclusion

Urban primacy, the extent to which a country's largest city dominates the national urban system, varies considerably across countries. Panama City comprises more than 70% of Panama's total urban population, while in neighboring Costa Rica only 42% of the national

urban population are concentrated in San José, and the ratio even drops to 28% for Honduras's largest city, Tegucigalpa.

A potential explanation for these differences that has recently gained considerable prominence is that differences in external openness might matter: countries open to international trade tend to have less dominant central cities than close economies, other things equal. In this paper, I examine the empirical evidence for this hypothesis. Providing a large variety of empirical tests, I find at best only weak support for the claim that trade liberalization reduces urban concentration.

References:

Ades, Alberto F. and Edward L. Glaeser. 1995. „Trade and Circuses: Explaining Urban Giants," Quarterly Journal of Economics. 110 (February): 195-227.

Frankel, Jeffrey A. and David Romer. 1999. „Does Trade Cause Growth?" American Economic Review. 89 (June): 379-399.

Hanson, Gordon H. 1998. „Regional Adjustment to Trade Liberalization," Regional Science and Urban Economics. 28 (July): 419-444.

Henderson, J. Vernon. 1996. "Ways to Think About Urban Concentration: Neoclassical Urban Systems Versus the New Economic Geography," International Regional Science Review. 19: 31-36.

Henderson, J. Vernon. 2000. „The Effects of Urban Concentration on Economic Growth," NBER Working Paper #7503.

Henderson, J. Vernon. 2002. „Urban Primacy, External Costs, and Quality of Life," Resource and Energy Economics. 24 (February): 95-106.

Henderson, J. Vernon. 2003. „The Urbanization Process and Economic Growth: The So-What Question," Journal of Economic Growth. 8 (March): 47-71.

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 Gordon H. Hanson. 1993. „Mexico-U.S. Free Trade and the Location of Production," in Peter M. Garber (ed.) The Mexico-U.S. Free Trade Agreement. Cambridge, MA: MIT Press.

Krugman, Paul R. and Raul Livas Elizondo. 1996. „Trade Policy and Thirld World Metropolis," Journal of Development Economics. 49 (April): 137-150.

Nitsch, Volker. 2001. City Growth in Europe. Berlin: Duncker & Humblot.

Rappaport, Jordan and Jeffrey D. Sachs. 2003. "The United States as a Coastal Nation," Journal of Economic Growth. 8 (March): 5-46.

Rodrik, Dani and Francisco Rodriguez. 2000. „Trade Policy and Economic Growth: A Skeptic's Guide to the Cross-National Evidence," in Ben Bernanke and Kenneth Rogoff (eds.) NBER Macroeconomics Annual 2000. Cambridge, MA: MIT Press.

Rose, Andrew K. 2003. „Do WTO Members Have More Liberal Trade Policy?" Journal of International Economics. (forthcoming).

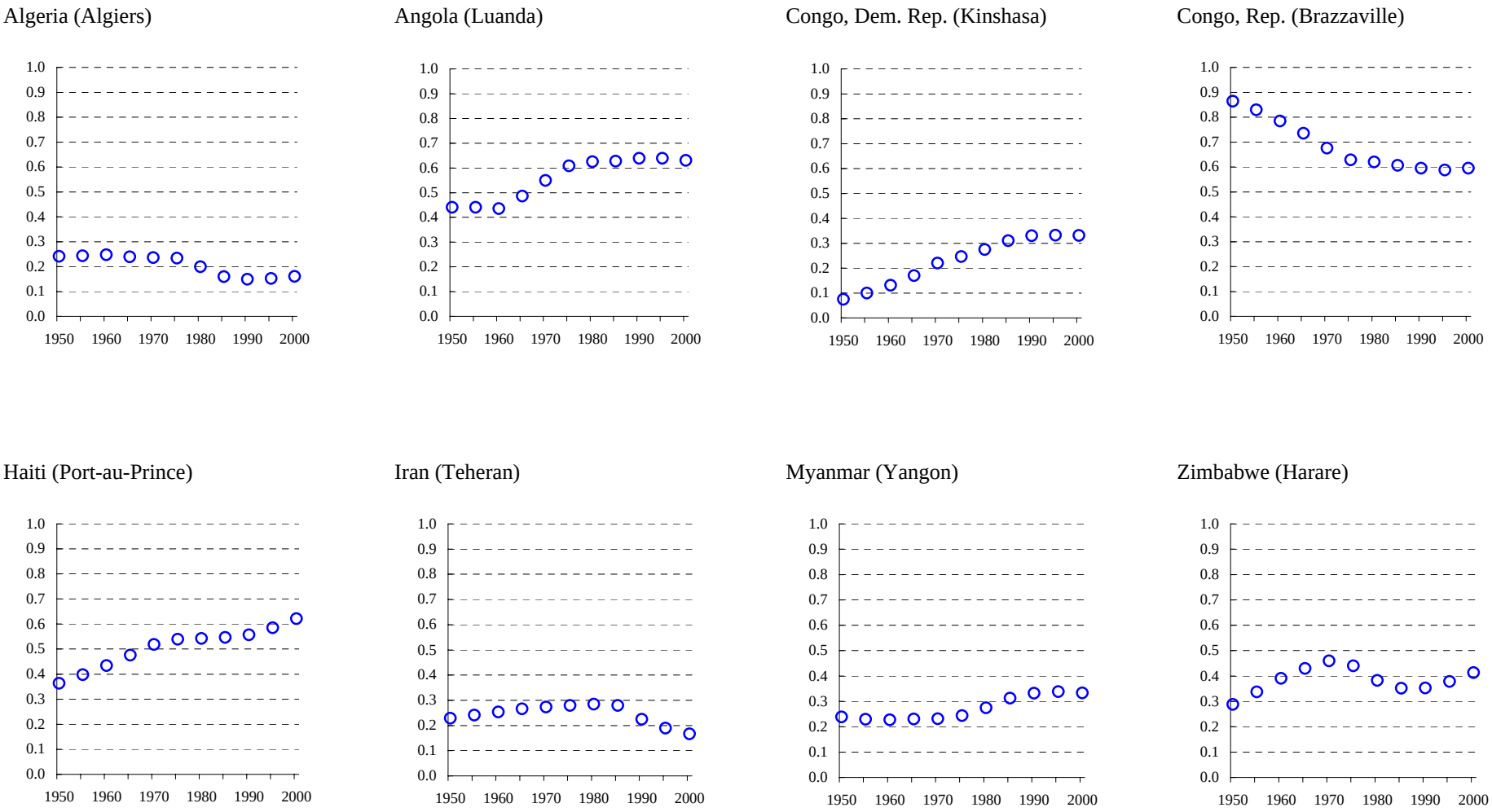
Sachs, Jeffrey D. and Andrew Warner. 1995. „Economic Reform and the Process of Global Integration," Brookings Papers on Economic Activity. (1): 1-118.

Venables, Anthony J. 2000. „Cities and Trade: External Trade and Internal Geography in Developing Economies," in Shahid Yusuf, Weiping Wu and Simon Evenett (eds.), Local Dynamics in an Era of Globalisation: 21st Century Catalysts for Development. Washington, DC: World Bank.

Wacziarg, Romain and Karen Horn Welch. 2002. „Trade Liberalization and Growth: New Evidence," Stanford University.

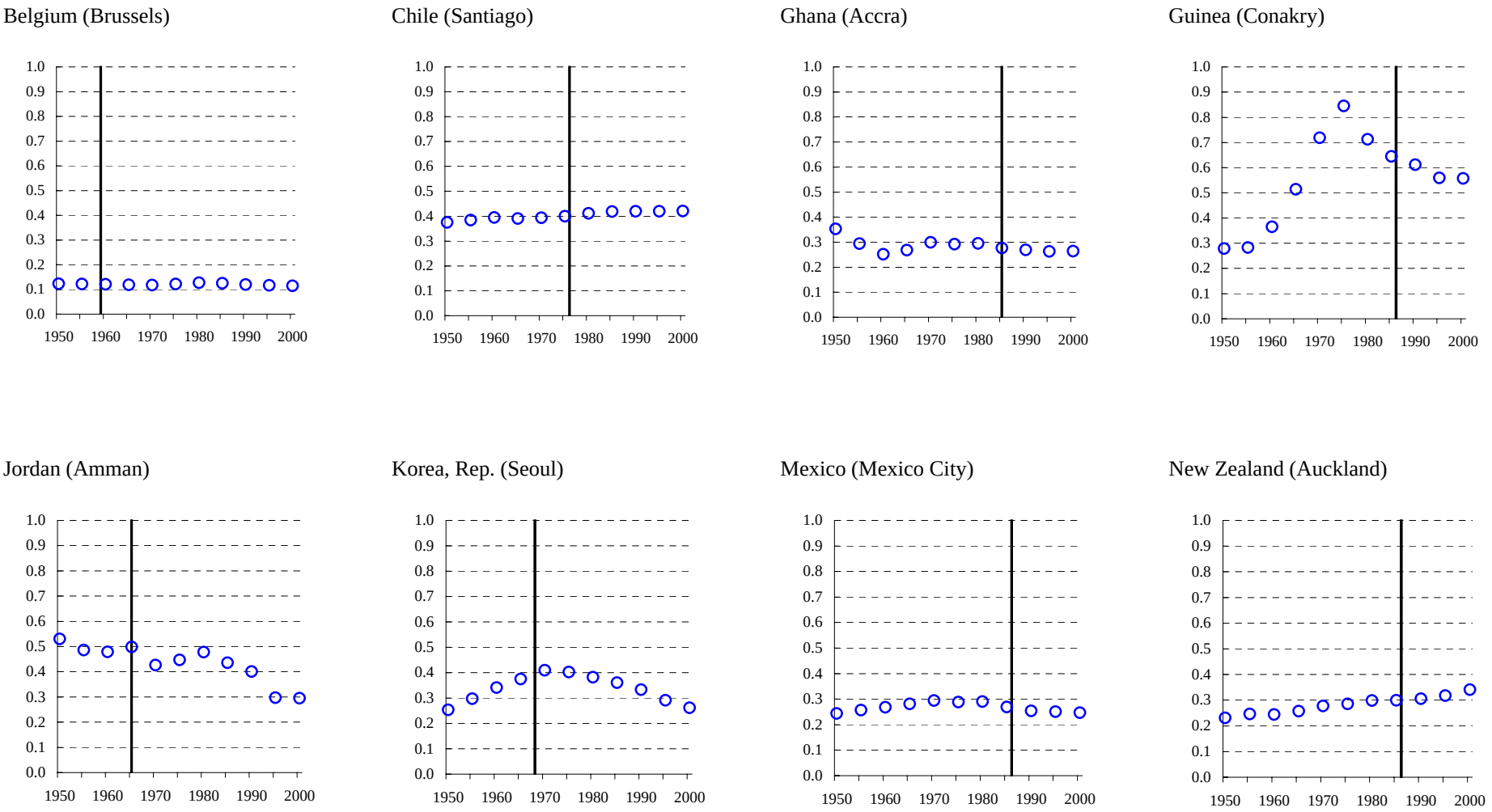
World Bank. 2003. World Development Report 2003: Sustainable Development in a Dynamic World. Washington, DC: World Bank.

Figure 1: Urban Primacy and Trade Liberalization — Closed Countries



Note: These countries were consistently classified as closed.

Figure 2: Urban Primacy and Trade Liberalization — Open Countries



Note: The vertical line shows the date of trade liberalization.

Table 1: Minor Modifications

	A&G (1995)	Basic	More countries	Other period	
Time period	1970-85	1970-85	1970-85	1985-2000	1985-2000
Country sample	A&G	A&G	Full sample	A&G	Full sample
Capital city dummy	0.465*	0.426#	0.509**	0.383*	0.480**
	(0.196)	(0.220)	(0.140)	(0.163)	(0.148)
Log of nonurbanized population	0.553**	0.521**	0.582**	0.364**	0.421**
	(0.066)	(0.090)	(0.059)	(0.076)	(0.066)
Log of urbanized population outside the main city	0.066	0.074	0.035	0.151*	0.130*
	(0.045)	(0.051)	(0.047)	(0.061)	(0.063)
Log of land area	0.155**	0.173**	0.085*	0.119*	0.090#
	(0.049)	(0.053)	(0.040)	(0.048)	(0.047)
Log of real GDP per capita	0.058	-0.098	0.295#	0.120	0.102
	(0.131)	(0.176)	(0.166)	(0.113)	(0.109)
Share of the labor force outside of agriculture	2.556**	3.274**	2.151**	1.837**	1.964**
	(0.567)	(0.646)	(0.603)	(0.480)	(0.450)
Share of trade in GDP	-0.609**	-0.547#	-0.682**	-0.586*	-0.722**
	(0.225)	(0.314)	(0.251)	(0.238)	(0.223)
Number of observations	85	84	115	79	108
Adjusted R^2	0.81	0.82	0.86	0.81	0.83

Notes: OLS estimation. Dependent variable is the log of population in the main city. The regressions are based on averaged data for the given period, available in five-year-intervals. White heteroskedastic-consistent standard errors are in parentheses. **, * and # denote significant at the 1%, 5% and 10% level, respectively. Constant not reported.

Table 2: Does Geographic Location Matter?

Time period	1970-85				
Capital city dummy	0.559**	0.558**	0.499**	0.494**	0.507**
	(0.144)	(0.144)	(0.152)	(0.153)	(0.151)
Log of nonurbanized population	0.590**	0.587**	0.580**	0.577**	0.577**
	(0.057)	(0.060)	(0.060)	(0.061)	(0.060)
Log of urbanized population outside the main city	0.026	0.029	0.036	0.038	0.032
	(0.046)	(0.048)	(0.048)	(0.050)	(0.050)
Log of land area	0.096*	0.100*	0.085*	0.087*	0.107**
	(0.040)	(0.041)	(0.041)	(0.040)	(0.041)
Log of real GDP per capita	0.294#	0.288#	0.294#	0.291#	0.283#
	(0.157)	(0.150)	(0.166)	(0.173)	(0.153)
Share of the labor force outside of agriculture	2.177**	2.192**	2.146**	2.151**	2.176**
	(0.570)	(0.557)	(0.610)	(0.622)	(0.564)
Share of trade in GDP	-0.696**	-0.795#	-0.680**	-0.645#	-0.709
	(0.242)	(0.424)	(0.251)	(0.327)	(0.589)
Port city dummy	0.199#	0.103			0.179
	(0.106)	(0.246)			(0.280)
Port city × Share of trade in GDP		0.165			0.121
		(0.434)			(0.515)
Central location dummy			0.021	0.068	0.191
			(0.108)	(0.232)	(0.249)
Central location × Share of trade in GDP				-0.077	-0.111
				(0.381)	(0.437)
Number of observations	115	115	115	115	115
Adjusted R^2	0.86	0.86	0.86	0.86	0.86

Notes: OLS estimation. Dependent variable is the log of population in the main city. The regressions are based on data for the given period available in five-year-intervals. White heteroskedastic-consistent standard errors are in parentheses. **, * and # denote significant at the 1%, 5% and 10% level, respectively. Constant not reported.

Table 2 (continued): Does Geographic Location Matter?

Time period	1985-2000				
Capital city dummy	0.513**	0.505**	0.489**	0.496**	0.483**
	(0.145)	(0.143)	(0.149)	(0.155)	(0.146)
Log of nonurbanized population	0.436**	0.424**	0.424**	0.427**	0.415**
	(0.068)	(0.069)	(0.069)	(0.070)	(0.068)
Log of urbanized population outside the main city	0.110	0.122#	0.129*	0.126#	0.126#
	(0.067)	(0.070)	(0.064)	(0.065)	(0.068)
Log of land area	0.096*	0.100*	0.087#	0.086#	0.107*
	(0.046)	(0.046)	(0.047)	(0.048)	(0.047)
Log of real GDP per capita	0.089	0.081	0.102	0.104	0.074
	(0.107)	(0.108)	(0.109)	(0.113)	(0.111)
Share of the labor force outside of agriculture	2.054**	2.071**	1.976**	1.969**	2.081**
	(0.445)	(0.454)	(0.455)	(0.460)	(0.465)
Share of trade in GDP	-0.761**	-0.888**	-0.730**	-0.771*	-0.798*
	(0.211)	(0.310)	(0.224)	(0.361)	(0.378)
Port city dummy	0.224*	0.057			0.077
	(0.106)	(0.242)			(0.251)
Port city × Share of trade in GDP		0.252			0.257
		(0.371)			(0.383)
Central location dummy			-0.038	-0.083	0.143
			(0.102)	(0.268)	(0.256)
Central location × Share of trade in GDP				0.065	-0.140
				(0.409)	(0.384)
Number of observations	108	108	108	108	108
Adjusted R^2	0.83	0.84	0.83	0.83	0.84

Notes: OLS estimation. Dependent variable is the log of population in the main city. The regressions are based on data for the given period available in five-year-intervals. White heteroskedastic-consistent standard errors are in parentheses. **, * and # denote significant at the 1%, 5% and 10% level, respectively. Constant not reported.

Table 3: Urban Primacy Measures as Dependent Variable

Dependent variable	Log of share of main city in urban pop. outside the main city				Log of share of main city in total urban population			
	Basic	More countries	Other period		Basic	More countries	Other period	
Time period	1970-85	1970-85	1985-2000	1985-2000	1970-85	1970-85	1985-2000	1985-2000
Country sample	A&G	Full sample	A&G	Full sample	A&G	Full sample	A&G	Full sample
Capital city dummy	0.593*	0.393#	0.489*	0.493**	0.352	0.309	0.388*	0.443**
	(0.237)	(0.202)	(0.208)	(0.167)	(0.260)	(0.186)	(0.173)	(0.150)
Log of nonurbanized population	-0.464**	-0.425**	-0.450**	-0.419**				
	(0.097)	(0.079)	(0.077)	(0.061)				
Log of total population					-0.650**	-0.559**	-0.609**	-0.537**
					(0.109)	(0.080)	(0.071)	(0.057)
Log of land area	0.158	0.067	0.123#	0.084	0.140	0.067	0.115	0.083
	(0.101)	(0.084)	(0.072)	(0.076)	(0.126)	(0.091)	(0.082)	(0.077)
Log of real GDP per capita	-0.500*	0.242	-0.077	-0.087	-0.418	0.106	0.029	-0.034
	(0.243)	(0.346)	(0.177)	(0.157)	(0.253)	(0.250)	(0.164)	(0.147)
Share of the labor force outside of agriculture	2.548*	-0.276	0.850	0.590	2.840**	0.829	0.920	0.928
	(0.958)	(1.325)	(0.834)	(0.684)	(1.010)	(0.918)	(0.751)	(0.642)
Share of trade in GDP	-1.137#	-0.410	-0.572#	-0.427	-1.699*	-1.003#	-0.898**	-0.811*
	(0.605)	(0.418)	(0.332)	(0.335)	(0.794)	(0.512)	(0.332)	(0.328)
Number of observations	84	115	78	107	84	115	79	108
Adjusted R^2	0.33	0.31	0.39	0.38	0.39	0.38	0.48	0.46

Notes: OLS estimation. A logistic transformation is applied when the share of the main city in total urban population is the dependent variable. The regressions are based on data for the given period available in five-year-intervals. White heteroskedastic-consistent standard errors are in parentheses. **, * and # denote significant at the 1%, 5% and 10% level, respectively. Constant not reported.

Table 4: Are Main Cities Different?

Time period City sample	1970-85 Main City	1970-85 Top 2 Cities	1970-85 Top 3 Cities	1970-85 Top 4 Cities	1970-85 Top 5 Cities	1970-85 Top 5 Cities
Capital city dummy	0.414** (0.119)	0.633** (0.116)	0.540** (0.113)	0.415** (0.142)	0.376** (0.136)	0.412** (0.130)
Capital city × Largest city		-0.187 (0.180)	-0.060 (0.184)	0.097 (0.218)	0.160 (0.224)	0.096 (0.207)
Log of nonurbanized population	0.559** (0.060)	0.577** (0.082)	0.542** (0.080)	0.539** (0.077)	0.532** (0.086)	0.552** (0.064)
Log of urbanized population outside the main city	0.052 (0.049)	0.069 (0.065)	0.137* (0.064)	0.143* (0.056)	0.150* (0.063)	0.107* (0.047)
Log of land area	0.074 (0.047)	0.130* (0.050)	0.144** (0.049)	0.166** (0.053)	0.186** (0.055)	0.161** (0.043)
Log of real GDP per capita	0.272 (0.175)	-0.010 (0.147)	0.116 (0.132)	0.132 (0.145)	0.148 (0.152)	0.094 (0.125)
Share of the labor force outside of agriculture	2.117** (0.657)	3.069** (0.568)	2.532** (0.504)	2.503** (0.543)	2.414** (0.565)	2.649** (0.469)
Share of trade in GDP	-0.746* (0.310)	-0.278 (0.417)	-0.078 (0.377)	0.009 (0.389)	0.178 (0.388)	0.005 (0.316)
Largest city × Share of trade in GDP		-0.025 (0.277)	-0.066 (0.268)	-0.053 (0.288)	-0.168 (0.282)	-0.378 (0.301)
Number of observations (countries)	103 (103)	158 (79)	219 (73)	272 (68)	320 (64)	394 (103)
Adjusted R^2	0.85	0.85	0.86	0.86	0.88	0.86

Notes: OLS estimation. Dependent variable is the log of city population. The regressions are based on data for the given period available in five-year-intervals. White heteroskedastic-consistent standard errors are in parentheses. **, * and # denote significant at the 1%, 5% and 10% level, respectively. Constant and controls for city rank in the national size distribution not reported.

Table 4 (continued): Are Main Cities Different?

Time period City sample	1985-2000 Main City	1985-2000 Top 2 Cities	1985-2000 Top 3 Cities	1985-2000 Top 4 Cities	1985-2000 Top 5 Cities	1985-2000 Top 5 Cities
Capital city dummy	0.350** (0.128)	0.669** (0.197)	0.571** (0.167)	0.500** (0.159)	0.473** (0.156)	0.499** (0.146)
Capital city × Largest city		-0.264 (0.256)	-0.173 (0.237)	-0.103 (0.242)	0.046 (0.243)	-0.061 (0.221)
Log of nonurbanized population	0.445** (0.065)	0.497** (0.089)	0.416** (0.084)	0.381** (0.085)	0.345** (0.090)	0.415** (0.072)
Log of urbanized population outside the main city	0.123# (0.062)	0.084 (0.086)	0.177* (0.076)	0.208** (0.077)	0.220** (0.077)	0.169** (0.063)
Log of land area	0.069 (0.048)	0.128* (0.050)	0.133** (0.049)	0.139** (0.051)	0.158** (0.050)	0.142** (0.039)
Log of real GDP per capita	0.086 (0.111)	0.011 (0.120)	0.038 (0.129)	0.003 (0.123)	0.056 (0.133)	0.037 (0.110)
Share of the labor force outside of agriculture	2.011** (0.471)	2.460** (0.513)	2.231** (0.541)	2.305** (0.514)	2.075** (0.572)	2.285** (0.451)
Share of trade in GDP	-0.612** (0.225)	-0.484# (0.255)	-0.436# (0.222)	-0.519* (0.251)	-0.569* (0.252)	-0.419* (0.194)
Largest city × Share of trade in GDP		0.180 (0.207)	0.158 (0.218)	0.098 (0.288)	-0.070 (0.295)	-0.061 (0.205)
Number of observations (countries)	111 (111)	174 (87)	237 (79)	300 (75)	340 (68)	424 (111)
Adjusted R^2	0.81	0.81	0.82	0.84	0.85	0.83

Notes: OLS estimation. Dependent variable is the log of city population. The regressions are based on data for the given period available in five-year-intervals. White heteroskedastic-consistent standard errors are in parentheses. **, * and # denote significant at the 1%, 5% and 10% level, respectively. Constant and controls for city rank in the national size distribution not reported.

Table 5: The Effects of Trade Liberalization on Urban Primacy

a) Liberalization date derived from year when uninterrupted openness began

Time period	1950-2000	1950-1975	1975-2000
	Country fixed effects		
Liberal trade regime	-0.010#	-0.005	-0.011#
	(0.005)	(0.005)	(0.006)
Number of observations	1,122	612	612
Adjusted R^2	0.91	0.94	0.95
	Country fixed effects with trend		
Liberal trade regime	-0.012#	-0.017**	-0.005
	(0.006)	(0.006)	(0.006)
Year	-0.0001	0.0008**	-0.0004#
	(0.0001)	(0.0003)	(0.0002)
Number of observations	1,122	612	612
Adjusted R^2	0.91	0.94	0.95
	Country and year fixed effects		
Liberal trade regime	-0.012#	-0.018**	-0.008
	(0.006)	(0.006)	(0.007)
Number of observations	1,122	612	612
Adjusted R^2	0.91	0.94	0.95

Notes: OLS estimation. Dependent variable is urban primacy defined as the share of the main city in total urban population. The regressions are based on data for the given period available in five-year-intervals. Huber-White robust standard errors are in parentheses. **, * and # denote significant at the 1%, 5% and 10% level, respectively. Country and (if applicable) year effects not reported. The sample comprises 102 countries.

Table 5 (continued): The Effects of Trade Liberalization on Urban Primacy

b) Additionally allowing for periods of temporary liberalization

Time period	1950-2000	1950-1975	1975-2000
	Country fixed effects		
Liberal trade regime	-0.008	-0.001	-0.010
	(0.005)	(0.004)	(0.006)
Number of observations	1,122	612	612
Adjusted R^2	0.91	0.94	0.95
	Country fixed effects with trend		
Liberal trade regime	-0.008	-0.007	-0.004
	(0.005)	(0.005)	(0.006)
Year	-0.0000	0.0007*	-0.0004#
	(0.0001)	(0.0003)	(0.0002)
Number of observations	1,122	612	612
Adjusted R^2	0.91	0.94	0.95
	Country and year fixed effects		
Liberal trade regime	-0.007	-0.007	-0.006
	(0.005)	(0.005)	(0.006)
Number of observations	1,122	612	612
Adjusted R^2	0.91	0.94	0.95

Notes: OLS estimation. Dependent variable is urban primacy defined as the share of the main city in total urban population. The regressions are based on data for the given period available in five-year-intervals. Huber-White robust standard errors are in parentheses. **, * and # denote significant at the 1%, 5% and 10% level, respectively. Country and (if applicable) year effects not reported. The sample comprises 102 countries.

Table 6: Taking Endogeneity Seriously

Dependent variable	Log of population in the main city		Log of share of main city in urban pop. outside main city		Log of share of main city in total urban population	
Time period	1970-85	1985-2000	1970-85	1985-2000	1970-85	1985-2000
Capital city dummy	0.501**	0.483**	0.389#	0.485**	0.291	0.428*
	(0.152)	(0.171)	(0.208)	(0.182)	(0.202)	(0.171)
Log of nonurbanized population	0.658**	0.484**	-0.368**	-0.378**		
	(0.060)	(0.075)	(0.077)	(0.066)		
Log of total population	0.037	0.127*			-0.454**	-0.469**
	(0.044)	(0.063)			(0.069)	(0.066)
Log of land area	0.113*	0.129*	0.102	0.123	0.082	0.108
	(0.051)	(0.060)	(0.096)	(0.089)	(0.105)	(0.095)
Log of real GDP per capita	0.285	0.195	0.234	-0.025	0.083	0.043
	(0.184)	(0.121)	(0.349)	(0.164)	(0.273)	(0.163)
Share of the labor force outside of agriculture	2.118**	1.502**	-0.293	0.367	0.680	0.482
	(0.669)	(0.489)	(1.331)	(0.707)	(1.022)	(0.704)
Constructed share of trade in GDP	0.0015	0.0004	0.0048	0.0012	-0.0032	-0.0034
	(0.0061)	(0.0062)	(0.0109)	(0.0096)	(0.0102)	(0.0081)
Number of observations	115	102	115	101	115	102
Adjusted R^2	0.84	0.81	0.31	0.36	0.35	0.41

Notes: IV estimation. A logistic transformation is applied when the share of the main city in total urban population is the dependent variable. The regressions are based on data for the given period available in five-year-intervals. White heteroskedastic-consistent standard errors are in parentheses. **, * and # denote significant at the 1%, 5% and 10% level, respectively. Constant not reported.