NBER WORKING PAPER SERIES

THE EFFECTS OF URBAN CONCENTRATION ON ECONOMIC GROWTH

Vernon Henderson

Working Paper 7503
http://www.nber.org/papers/w7503

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2000

Support of the World Bank through a research grant is gratefully acknowledged , as well as support of the National Science Foundation. The opinions and conclusions in this paper are those of the author and do not necessarily represent those of the World Bank, the National Bureau of Economic Research or NSF. I thank Zmarak Shalizi for many helpful discussions. James Davis excelled as a RA on the project. I benefited from advice from Andrew Foster and David Weil on various aspects of the paper and I benefited in conceptualizing the problem from early discussions with Shahid Yusuf. I benefited from comments on the paper by Duncan Black, Jim Davis, Yannis Ioannides, and Will Strange.

The Effects of Urban Concentration on Economic Growth
Vernon Henderson
NBER Working Paper No. 7503
January 2000
JEL No. O1, O4, O57, R1

ABSTRACT

The paper examines whether there is a significant relationship between economic growth and the degree of urban concentration, as measured by primacy, or the share of the largest metro area in national urban population. Is there reason to believe many countries have excessive primacy and how costly is excessive (or insufficient) primacy? Using GMM methods, the paper estimates growth effects, using a panel of 80-100 countries from 1960 to 1995. It also looks at the determinants of primacy and policy instruments that might be effective in reducing excessive primacy. The paper finds that there is a best degree of national urban primacy, which increases sharply up to a per capita income of about $5000 (PPP 1987 income), before declining modestly. The best degree of primacy declines with country scale. Error bands about estimated best degrees of primacy are generally tight. Growth losses from significantly non-optimal concentration are large and rise with income. Results are very robust. In a group of 72 countries in 1990, it appears that at least 24 have satisfactory primacy; at least 24 have significantly excessive primacy; and at least 5 countries have too little.

What determines urban concentration? Econometric models show that urban concentration initially rises with income and then peaks around an income of $2400, before declining. Openness, or trade effects are modest. Similarly, the effects of a greater degree of political decentralization while significantly reducing urban concentration are quite modest. The key policy type variable affecting concentration is investment in inter-regional transport infrastructure. In particular, increases in the density of road networks significantly reduce primacy, with the effect rising with income. As a policy consideration, this takes heightened importance because growth losses from excessive primacy tend to rise with income. The effect on growth rates of investment in roads, through its effect on primacy, is highest in middle income countries.

J. Vernon Henderson
Department of Economics
Box B
Brown University
Providence, RI 02912
and NBER
j_henderson@brown.edu

Urbanization and economic growth in developing countries go hand-in-hand. The simple correlation coefficient across countries between the percent urbanized in a country and GDP per capita (in logs) is about 0.85. The reason is clear. Economic development involves the transformation of a country from an agricultural based economy to an industrial-service based economy. Production of manufacturing and services is much more efficient when concentrated in dense business-industrial districts in cities. Close spatial proximity, or high density, promotes information spillovers amongst producers, more efficiently functioning labor markets, and savings in the transport costs of parts and components exchange among producers and of sales to local residents. The existence and considerable magnitude of localized scale externalities is well documented empirically (Henderson (1988), Ciccone and Hall (1995), Glaeser et al. (1992)). The transport savings component of high density is central to the new economic geography literature (Fujita, Krugman, and Venables (1999)) and is starting to be documented, especially for face-to-face transactions costs in services (Kolko (1999)).

While national policies can retard or accelerate urbanization rates, the concern in this paper is not with urbanization per se, but rather with the form urbanization takes. By form, I mean the degree of urban concentration. At any point in time, given a country's level of urbanization, resources may be spread too thinly/evenly across cities with insufficient concentration in certain cities to exploit the economies of scale in production which were cited above. Alternatively, resources may be over-concentrated in one or two excessively large cities, raising commuting, congestion, and living costs to excessive levels, raising costs of production of goods and lowering the quality of urban service provision. The implication is that there is an optimal degree of urban concentration, achieved by trading-off the social marginal benefits and costs of increasing urban concentration. Either over or under-concentration, as we will see, is very costly in terms of economic efficiency and national growth rates.

There is also a dynamic component to this discussion of optimal urban concentration. In the development literature (Williamson (1965)), as adapted to an urban context in Hansen (1990)), a high degree of urban concentration in the early stages of economic development is viewed as essential to efficiency. By spatially concentrating industrialization, often in coastal cities, the economy conserves on "economic infrastructure" – physical infrastructure capital (transport and telecommunications) and managerial resources. Such spatial concentration also enhances information spillovers at a time when the economy is "information deficient" and it may similarly enhance knowledge accumulation (Lucas (1988), Black and Henderson (1999)). As development proceeds, eventually deconcentration becomes efficient for two reasons. The economy can afford to spread economic infrastructure and knowledge resources to hinterland areas. Second, the cities of initial high concentration become high cost, congested locations that are less efficient locations for producers and consumers.[1] On the positive, empirical side, Wheaton and Shishido (1981) find the pattern of first increasing and then decreasing urban concentration across countries as income rises, a result consistent with findings of regional convergence within countries over time.

Whatever the best degree of urban concentration at any point in time, there is a presumption in the literature, both that countries have a tendency to urban over-concentration, and that urban over-concentration is costly to economic growth. The Williamson initial concentration process proceeds too far and deconcentration is delayed too long. There are three strands to the literature. The theoretical urban literature argues that, at any point in time, the various city sizes across the economy will only be efficient if national land development markets work perfectly (Henderson (1974), Becker and Henderson (1999)). That perfect working requires new cities to be able to form freely, usually through the initiative of large-scale land developers and local governments, in a context where there are strong institutions governing land markets and contracts and there is complete local fiscal autonomy. If these conditions are not in place, cities will be oversized, resulting in excessive urban concentration (as in a self-organized world (Fujita et al. (1999)) where no developers/local governments act to aid the development of cities). In all models, having under-sized cities does not constitute a stable equilibrium; cities are either efficient-sized or over-sized.[2] Caveats are that efficient city size is defined for a given level of technology and amenities and cities are assumed to efficiently subsidize industrial location so as to internalize local scale externalities. If

[1] Deconcentration occurs by manufacturing moving first from the core cities of large metro areas to nearby satellite cities, and then into hinterland cities, where wage and land costs are much lower. The initial large metro areas typically switch into more service-oriented production (financial, business, engineering and management, education, and health services). As the satellite cities and hinterlands industrialize, they increasingly become the locations of choice for migrants. Growth rates of the very largest cities tend to slow, while those of medium and large size cities continue unabated. Indeed in developing countries worldwide, the population growth rate of metro areas over 5m is a half that of metro areas in the 0.5m range (WDR 2000, Table 6.3).

[2] Any city of the same type which is a little larger than an under-sized one will attract resources from the under-sized city, since it is not exploiting scale economies sufficiently.

technology is relatively deficient in a city or local externalities are not fully internalized, a city can be too small relative to its efficient size in a better equilibrium.

The second strand of the literature tries to assess the costs and benefits of increasing city sizes (Tolley, Gardner, and Graves (1979) and Richardson (1987)). Tolley et al.'s empirical work tends to suggest that in large cities the social marginal costs of increasing city population exceed the marginal benefits and such cities are over-sized. Richardson (1987) focuses on the investment cost side. Based on work on Bangladesh, Egypt, Indonesia, and Pakistan, he argues that the social investment costs of absorbing an extra family in typical large urban areas are threefold that of rural areas, and even more for the largest city in a country. As related evidence on private costs, UN data for 1996 on metro area rents and commuting costs for a sample of 80-100 cities across 15-20 countries world-wide suggest that, if urban area sizes increase from 25,000 to 2.5 million, commuting and rent costs each rise by 115% (Henderson (1999b)). This literature tends to presume that the social marginal benefits of moving a family to large urban areas don't justify the various costs, implying these cities are over-sized.

The final strand of literature (Renaud (1981), Henderson (1988), Ades and Glaeser (1995)) argues that often the political institutions in countries encourage over-concentration. The idea is that, in many countries, there is a lack of a level playing field across cities. The national government can choose to favor one (or two) cities over others. Typically such cities are national capitals (Bangkok, Mexico City, Jakarta, or Seoul, not to mention Paris); but they may also be a Sao Paulo, the seat of national elites. Such favoritism can involve the allocation of local public services in favor of national capitals, where decision-makers live. That problem can be exacerbated if hinterland cities do not have the power to determine their own public service levels, either because of a unitary national constitution or because local autonomy has been suspended (as in Korea from 1961 to the 1990's).

Favoritism can take the form of the national government choosing not to invest sufficiently in interregional transport and telecommunications, so that hinterland cities are less competitive locations for private producers. That favors producers and investors (who may include national politicians) in the national capital. Favoritism, as in Indonesia (Henderson and Kuncoro (1996) and Kaiser (1999)), can also take the form of restrictions in capital markets, export/import markets, and licensing of production rights, all favoring firms which locate in the national capital. This allows central bureaucrats and politicians to extract rents in the allocation of loans and licenses, without competition from lower ranked bureaucrats in other locations. Finally, there may be an “innocent” bias towards locating production in the city that national decision- makers are most informed about – perhaps they believe such constraints are efficient.

All analyses tell us favored cities are oversized with attendant efficiency losses for the country. Migrants and firms flow to a favored city, until it becomes so congested and costly to live in, that these costs offset the advantages of the favoritism. Moreover, the excessive resources devoted to one or two favored cities detract from the quality of life in the rest of the urban system. In Henderson (1999b), based on the UN data set for 80-100 cities in 1996 worldwide, I show that high urban concentration in a nation increases child mortality, pupil-teacher ratios, use of non-potable water and other poor quality of life dimensions in typical medium size metro areas, after accounting for size, income, and growth differences among cities. For example, a one-standard deviation increase in the national urban concentration measure raises child mortality in typical cities by 1/3 of a standard deviation of the child mortality rate across cities in the sample. So the costs of excessive urban concentration are felt throughout the whole urban system, not just in very large cities.

Apart from analyses and empirical work in economics, international agencies also take the view that many of the world’s mega-cities are over-populated, at considerable cost to those economies. The UN (1993) asks how bad “the negative factors associated with very large cities” need to get “before [it is in the] self-interest of those in control to encourage development of alternative centers.” The same report warns of “unbalanced urban hierarchies” and the crime, congestion, and social inequality in mega-cities. Some World Bank work (Renaud (1981)) echoes these concerns and articulates the over-size bias. Even in the US, the popular press has noted the competitiveness of medium-size metro areas for most economic activity, compared to the very largest metro areas (1-5-99 N.Y. Times, Business section, p. 1).

This paper has three objectives:

1) Using data on economic growth and urban-concentration in five-year intervals from 1960-1995 for samples of 80-100 countries, the paper will examine the relationship between urban concentration and economic growth, at different stages of economic development in different size countries. **If urban concentration really is an issue, then it ought to affect economic growth rates in a robust, consistent fashion.** The basic notions are:

a) In any economy there are initial gains in economic growth rates from increasing concentration from low levels, but these gains peak, and further increases in urban concentration bring losses. That is, there is, hypothetically, a best degree of urban concentration.

b) Following Williamson-Hansen, the best degree of urban concentration initially increases as a country starts to grow from very low income levels. But, then, with further growth the desired degree of urban concentration declines.

c) The desired degree of urban concentration in general declines with country size.

d) While many countries seem to operate beyond the point of desired concentration, some do not.

2) The paper will examine the determinants of urban concentration in a country. The degree of urban concentration is determined by the stage of development and the country's size, as well as by institutions and national political processes. I will explore the role of institutions, in particular the degree of federation, or political decentralization in the country. The political economy literature suggests that increased regional representation and autonomy decreases centralization (WDR 2000). I will also explore the effect of policy outcomes, such as degree of openness to trade, investment in inter-regional infrastructure, and the role of national capitals. Investment in interregional infrastructure may decrease urban concentration, because hinterland areas become more competitive.

3) If urban concentration in many countries appears excessive and, for those countries, a deterrent to growth, what is the policy prescription? Of course, there are simple economic reforms that level the playing field across firms, people, and space, which would directly and indirectly (through reducing urban concentration) help growth. Quite apart from general economic liberalization, the paper will argue that investment in inter-regional transport infrastructure is a key policy instrument which, at the appropriate stage of economic development, would directly reduce concentration and thus indirectly substantially help spur economic growth.

In implementing this study, a key issue concerns how to measure urban concentration. While it would be desirable to use a Hirschman-Herfindahl type index based on the sum of squared shares of every city in a country in national urban population, such data are not available over the time period. What is available and what is utilized (e.g., Ades and Glaeser (1995)) is urban primacy – measured here as the share of the largest city in national urban population. While this could be a crude measure, because such shares are typically very large, primary measures tend to be closely correlated with Hirschman-Herfindahl indices (Henderson (1999a)). Since Hirschman-Herfindahl indices contain squared shares, they tend to be dominated by the largest share if that is a high number (e.g., 0.25). Average primacy in our sample, over countries and years is .30. This idea of close correlation is also supported by evidence on Zipf's Law (Gabaix (1999)). Within countries when we rank cities from largest (rank 1) to smallest, rank times population size is approximately the same constant for all cities. Thus the size of the largest city in the country defines all other city sizes and is sufficient information to calculate any comparative index of national urban concentration.

The issue I see in using urban primacy is not one of measurement, but rather, that the growth results in this paper could be used to proscribe city sizes. There is no question that, later in the paper, when I talk about a best degree of urban primacy (i.e., concentration) in a country of a given size and income level, that such a number could be translated into a population number for the largest city. However, that is not the point of this paper, per se. The paper is using primacy just as a general concentration index. But in thinking about specific primate city sizes, there will be error bands on the best size, which give a range of acceptable city sizes. These error bands reflect the fact that optimal city size depends on both measured and unmeasured specifics of a particular city and country– its industrial composition (or “type” in a national or international hierarchy), its location, and its level of knowledge or technology. Rather than proscribing city sizes, I identify outliers, or countries where primate cities lie outside the error bands. And the preferred solution in problem countries is to free markets and to establish institutional frameworks, so that market city sizes can approach efficient ones, rather than trying to set city sizes per se. In distorted situations with poor institutions, *unless* these distortions and institutions are improved to what might reasonably be expected for a country's level of economic development, the econometrics don't really tell us what is the best primate city size. They just tell us something is wrong in the operation of urban development markets, causing excessive concentration and growth losses.

1. **Urban Concentration and Economic Growth**

To incorporate considerations of urban concentration into an economic growth framework, I first review the derivation of a standard empirical growth model (see Durlauf and Quah (1998) for a nice synthesis of various theoretical and empirical approaches). Output in an economy is usually specified as produced according to an aggregate Cobb-Douglas function of the form

$$Y = (K(t))^{\alpha} (A(t) N(t))^{1-\alpha} \quad (1)$$

where K(t) is capital, N(t) is labor (proportional to population, with the factor of proportionality normalized to 1), and A(t) the level of technology. Labor and technology grow at rates n and g, so $\dot{N}/N = n$, $\dot{A}/A = g$. Capital depreciates at the rate δ, and s is the fraction of output saved and invested in a Solow model. Finally, output and capital per effective unit of labor are defined respectively as $\hat{y} = Y/AN$ and $\hat{k} = K/AN$. The steady state value of $\hat{k}$ is $\hat{k}^*$ (= $(s/(n+g+\delta))^{1/(1-\alpha)}$), where $\hat{y}^* = \hat{k}^{*\alpha}$.

To derive an estimating equation, we can do a linear expansion in natural logs of the equation of motion ($d\hat{k}/dt = s\hat{k}^{\alpha} - (n+g+\delta)\hat{k}$) about its steady state value, or we can solve the equation of motion for $\hat{y}$ as a function of $\hat{y}^*$ and time and do a linear expansion in logarithms of that expression for two time periods. Then combining for these two time periods, t_2 and t_1, we get

$$\log \hat{y}(t_2) - \log \hat{y}(t_1) = (1-e^{-\beta\tau}) \log \hat{y}^* - (1-e^{-\beta\tau}) \log \hat{y}(t_1) \quad (2)$$

where $\beta = (1-\alpha)(n+g+\delta)$ and $\tau = t_2 - t_1 > 0$.

β is the rate of convergence to the steady state. To convert (2) to observable magnitudes, we substitute in for $\hat{y}^*$ and for $\hat{y}(t) = y(t)/A(t)$. The result is

$$\log y(t_2) - \log y(t_1) = -(1-e^{-\beta\tau}) \log y(t_1) + (1-e^{-\beta\tau}) (\alpha/1-\alpha)$$
$$\log (s/(\delta+n+g)) + (1-e^{-\beta\tau}) \log A(t_1) + g\,\tau . \quad (3)$$

$A(t_1)$ is the level of technology. It is more common to write the last two terms as $(1-e^{-\beta\tau})\,A(0) + g\,(t_2 - e^{-\beta\tau}\,t_1)$. The expression $(\alpha/1-\alpha) \log (s/(\delta+n+g))$ can be replaced by a function $f(s/(\delta+n+g))$ if s is determined in a Cass (1965) optimization framework, rather than set exogenously.

Early empirical work calculated or assumed values of s, δ, n, and g for each country and estimated equation (3) directly, usually subsuming $g\,\tau\ [g(t_2 - e^{-\beta\tau} t_1)]$ into the constant term and allowing $\log A(t_1)$ $[\log A(0)]$ to be part of the error structure. Later work acknowledges that (a) RHS variables may not be exogenous to $A(t_1)$, (b) magnitudes such as n or s may only be able to be proxied, (c) δ or g may vary across countries, and (d) the exact representation in equation (3) assumes Solow savings behavior and approximations about steady-state values. Later work also incorporates the idea that internal country institutions and government policies may affect efficiency of the production process. So Barro (1991, 1997) allows inflation rates and government consumption to inhibit growth. Others have considered a variety of additional control variables such as civil liberties, exchange rate distortions, and availability of domestic credit. How to incorporate such considerations into (3) is not explicitly modeled. They are viewed generally as affecting $A(t_1)$ the base technology and level of efficiency, or g the rate of effective technological advance.

Based on these considerations, in more recent empirical approaches, formulations typically reduce for country i to

$$\log\ y_i(t_2) - \log y_i(t_1) = -\,(1-e^{-\beta\tau}) \log y_i(t_1) + X_i\,(t_1)\,\gamma + f_i + \eta_{t2} + \varepsilon_{it2} \qquad (4)$$

The $X_i(t_1)$ are a vector of determinants of country growth rates. My basic growth equation will specify the $X_i(t_1)$ to include the basics in Barro (1991) or Mankiw, Romer and Weil (1992), including the average investment rate from t_1 to t_2, the average fertility rate from t_1 to t_2, and a measure of human capital in t_1 -- the average number of years of high school and college education of the adult population (male and female). In addition, there will be an expanded model with the average rate of inflation and government consumption from t_1 to t_2. Note each five-year time interval in our data for each country will use the investment, fertility, and other rates for that five years (in essence, $\hat{y}^*$ is a shifting target).

To incorporate urban concentration, I assume that the form which urbanization takes affects efficiency of the country's production process. Inefficient urbanization reduces income levels and growth rates. Attempting to derive explicit growth equations to model inefficient urbanization in a Black and Henderson (1999) urban-growth framework where urbanization itself is efficient is an impossible task. Its impossible because (a) to model the determinants of why there is inefficiency (due to institutional arrangements and government restrictions) appears for now to be impossibly complex and (b) to capture the exact impacts of these determinants would involve very arbitrary specifications. However, we can use the framework in (4) to assess in reduced form the effect of certain characteristics of the urbanization process on growth itself. In modeling the effect of the urban process on growth, I will use primacy as the only available measure of concentration in the urban system. I tried a variety of experiments, based on preliminary results that increases in urban concentration alone are harmful to growth. These experiments are detailed below but they have the general form

$$p\,(\text{primacy}_i(t_1),\ \log y_i(t_1),\ \text{scale}_i(t_1)) \qquad (5)$$

where the arguments in $p(.)$ are respectively primacy, income, and country scale.

Equation (4) has an error structure of unmeasured attributes affecting growth. η_{t2} are time dummies representing global shocks, such as global technological advances. f_i are time invariant country characteristics representing geography and culture, and the ε_{it2} are contemporaneous shocks. Since the data are panel in nature we will be able to deal with the fact that the f_i affect the $X_i\,(t_1)$ and that the ε_{it2} may affect some of the $X_i\,(t_1)$ in the contemporaneous time period as well as in future time periods. Generally, we will present OLS, fixed effect, and instrumental variable (GMM) estimates, as detailed below.

Data

The data cover 1960-95 in five-year intervals (i.e., $\tau = 5$). Data on constant dollar income per capita (Chain index), investment share of GDP, and government consumption share of GDP (without national defense and education netted out) are from the Penn World Tables Mark 5.6. Data on inflation rates and total fertility rate (children per women) are from the World Bank's World Development Indicators [WDI]. Data on average years of high school and college education of the adult (over 25) population are from Barro and Lee (1996). Population data on total population, urban population and primacy (population of the largest metro area/national urban population) are from the UN World Urbanization Prospects, Tables A12, A.5 and A3. Other data sources will be cited as data are introduced and I footnote all relevant sources here.[3] Means and standard deviations of all variables are given in Table A3 in the Appendix.

For growth between t_1 and t_2 (e.g., 1990 and 1995), for investment share, fertility rate, inflation, and government consumption, the X_{it1} are the annual average rates over t_2-1 to t_1 (e.g., 1990-94). The Penn World Tables only go to 1992. Missing data to 1994 or 1995 (including 1995 income) are filled in using the WDI numbers. For example, 1995 income per capita is projected by growing the Penn World Tables constant 1992 income by WDI numbers on annual income growth from 1992-1995. Amongst missing observations are data on Czechoslovakia, Yugoslavia, USSR, and West Germany for 1995, with 1960-90 data defined for these country entities as they existed in 1990.

Econometric Implementation

OLS estimation of equation (4) (augmented by 5 when appropriate) pools all country-years and allows for time fixed effects (η_{t2}). Country fixed effect estimates control for the f_i. The remaining problem is that the contemporaneous error term ε_{it2} affecting growth from t_1 to t_2 may be correlated with investment and fertility rates from t_1 to t_2, although all other base period variables in equation (4) are considered exogenous to that growth period (i.e., predetermined). However there may be cross period correlation so that base period variables such as income, education, primacy or scale may be correlated with the ε_{it1} from the prior growth period. To deal with these problems I instrument in GMM estimation of differenced level equations (4).

Specifically, the estimating equation for any year, assuming a sequence of periods t_0, t_1, t_3, … is

[3] Barro, R.J., and J.W.Lee, International Measures of Schooling Years and Schooling Quality online data, World Bank Economic Growth Research Group, Washington D.C.: World Bank, 1996;
Central Intelligence Agency (CIA), World Factbook, Washington D.C.: US Government Printing Office, various years;
Freedom House, Freedom in the World, New York: Freedom House, various years.;
International Road Federation (IRF), World Road Statistics, Washington D.C.: International Road Federation, various years;
Summers, R., and A. Heston, Penn World Table Mark 5.6 revision of Summers and Heston (1991) online data, Computing in the Humanities and Social Sciences (CHASS), Toronto: University of Toronto, 1995;
United Nations, World Urbanization Prospects: The 1996 Revision, United Nations Population Division, Department for Economic and Social Information and Policy Analysis, New York, 1996;
and
World Bank, World Development Indicators (WDI) on CD-Rom, Washington D.C.: World Bank, 1998.

$$(\log y_i (t_3) - \log y_i (t_2)) - (\log y_i (t_2) - \log y_i (t_1)) = -(1-e^{-\beta\tau})(\log y_i (t_2) - \log y_i (t_1)) + (X_i (t_2) - X_i (t_1))\gamma + (\eta_{t3} - \eta_{t2}) + (\varepsilon_{it3} - \varepsilon_{it2}) \quad (6)$$

The fixed effects are differenced out. For X's which are in the form of rate variables going from t_2 to t_3-1 for example, ε_{it3} may be correlated with X_i (t_2). For base period variables, given differencing, ε_{it2} should be correlated with X_i (t_2). We instrument with predetermined values of variables (e.g., for rate variables, the average rate from t_0 to $t_1 - 1$, t_{-1} to $t_0 - 1$ and so on and for base period variables, values for t_1, t_0, t_{-1} and so on back). I also experimented with other instrumenting strategies and specifications, getting similar results.[4] The ε_{it} are assumed to be serially uncorrelated, although by differencing, the error terms in (6) are serially correlated between any two adjacent time periods.

In estimation, the number of equation-years is T_i-2, where T_i is the length of the panel (which may vary by country). One year is lost in differencing and one in instrumenting. Equation years are pooled constraining slope coefficients across years to be the same, accounting for year-to-year serial correlation. Unbalanced panels estimated in DPD98 (Arellano and Bond (1991)) are utilized. The two-step estimation procedure utilizes a (within year) heteroskedastic consistent estimate of the covariance matrix of moments. Instruments are all predetermined values of right-hand side variables, when DPD will accommodate them. In models with many right-hand side variables, sometimes I trim the instrument list, but the minimum is two periods of predetermined values. The assumptions on serial correlation are tested and hold (strongly) in all estimations (there is first-order serial correlation in differences, but none in levels, so there is no second-order correlation in differences). Sargan

[4] Why should predetermined values as instruments tell us about current changes in variables, outside the error structure of the estimating equation? There must be other processes in the economy not directly affecting growth that drive both predetermined variables and their current changes. Here for example, specifics of demographic structures and schooling enrollment rates of females and males help determine levels and changes in savings and fertility rates. Similarly items that affect primacy but not directly growth (see next section), such as the degree of political decentralization or specific infrastructure configurations help determine both changes in and past levels of primacy. I experimented with instrumenting with determinants of education levels, investment and primacy (see next section of the paper). These included growth in the young and old populations, growth in female and male high school and higher education enrollment rates, change in labor force participation, population growth, change in openness, and change in road density, while restricting other instruments to be only the exogenous (predetermined) variables in eq. (6), for just that equation year. Results are very similar to what I obtain below. Finally I estimated the equations in level form, using predetermined changes in RHS variables as instruments. Results with this approach are weaker, but the sign patterns of coefficients in the complex expressions below are the same, with similar implications. Estimation results are extremely robust.

tests on overidentifying restrictions are conducted, in part to assess the validity (exogeneity to the equation year) of the instrument lists.

Results for Basic Growth Formulation

I first examine the basic growth model without primacy consideration. I have a stripped down version consistent with Mankiw, Romer, and Weil (1992), Islam (1995) and Caselli, Esquivel, and Lefort (1996). Mankiw, Romer and Weil and then a more expanded version such as Barro (1997) utilizes. Since the data set is very extensive, and since the estimation technique accounts for endogeneity problems that are often ignored or poorly accounted for in the literature, I comment on fundamental growth issues. Results for the two specifications are given in Table 1, for the three estimation methods.

Results are given in Table 1. I start with a discussion of column 1 results, which don't include the Barro variables. My discussion focuses on the OLS and GMM results, with fixed effect results reported for those who are interested. Note the dependent variable is in five-year intervals. The coefficients on base period GDP per capita give the usual pattern (see Caselli et al. (1996)), with slow speeds of convergence (1.4%) for OLS, but much faster speeds under GMM (5.7%), or fixed effects. For other variables and coefficients, I follow the typical procedure of assessing the effect on growth rates of changes in variables. These effects could be more strictly interpreted as effects on steady state income levels with respect to changes in the associated variables. This is done by dividing by the convergence factor, the coefficient on base period income.

The annual average rate of investment raises growth. A one-standard deviation (8) increase in the investment rate for a coefficient of .007 raises average annual growth rates by about 1% point. Fertility has a negative effect but its magnitude and significance is sensitive to specification. But again, a one-standard deviation (.5) increase in its magnitude lowers growth rates by 1% also, for a coefficient of 0.1. For average years of schooling, OLS results (and fixed effect ones) are insignificant. But under GMM, results are significant and very strong. A one-standard deviation (1.1) increase in average years of schooling raises annual growth rates by 1.5%, for a coefficient of .07.

It appears GMM and OLS results are different for key variables. Speeds of convergence are much higher under GMM and knowledge accumulation effects are made stronger. Given the difference GMM makes and the strong case for using it, I tend to rely on those results. Below I report OLS and fixed effect results only for "best" specifications.

For Barro's policy variables, in column 2 results government consumption rates and inflation hurt growth, as he finds. A one-standard deviation (7) increase in the government consumption rate lowers annual growth by 1.4% under GMM but only 0.3% under OLS. For inflation, a one-standard deviation increase reduces growth by .2%, for a coefficient of .00007. The standard results are sensitive to inclusion of Barro's variables – convergence speed rises noticeably. Inclusion of these variables also reduces sample sizes (missing observations). Sometimes people insert the degree of openness of the economy, treating it as a policy rather than outcome variable. Here, openness (see later) produces a significant positive and large effect (with little impact on other coefficients and none on the primacy specification below). A one standard deviation increase in the degree of openness raises annual growth rates by over 3 %. In the primacy results to follow, which have complex interactions, to conserve sample size, I report results without these variables. However, I will note the basic effect of including these extra variables.

Effects of Urban Concentration on Growth

In estimation urban concentration has a complex relationship to growth, as anticipated in the more conceptual literature. As expected from the urban agglomeration literature, too little urban concentration is bad, as is too much concentration, so there is a best degree of urban concentration. However, as Williamson-Hansen anticipate, what is too little or too much changes with income. Initially, the best urban concentration point rises from low income. But then at some higher income level, the best degree peaks and then starts to decline with further income increases. Finally, we expect the best degree of urban concentration to decline with country size, other things being equal.

In terms of functional representation of the effect of urban concentration, or primacy interacted with income and country size, this is beyond a second-order Taylor series expansion. To limit terms and yet to capture what the literature suggests, I include only essential terms. I do three versions of equation (5), in increasing degree of complexity. The results for all these versions are robust to precise choice of instrumental variables, other independent variables, measures of national scale, sample of countries, etc. This robustness and the precision of results took me by surprise. They convinced me that these results are "very real": the form of urbanization has profound effects on economic growth rates.

First is to show that the effects of urban concentration vary with economic development, to (4) I add terms

$$+ \text{primacy}_i(t_1)\,[\delta_0 + \delta_1 \log y_i(t_1) + \delta_2 (\log y_i(t_1))^2]. \qquad (5a)$$

In (5a), I expect $\delta_1 > 0$ and $\delta_2 < 0$, so that the positive effects of primacy initially increase with income, up to a certain income level. Then with further increases in income, primacy becomes increasingly harmful. This is the basic Williamson hypothesis. A more sophisticated version recognizes the basic urban economics result that there can be too little or too much urban concentration at any income level. Then in (5b) we have

$$\begin{aligned} &+ \text{primacy}_i\,(t_i)\,[\delta_0 + \delta_1 \log y_i\,(t_1) + \delta_2 (\log y_i\,(t_1))^2] \\ &+ (\text{primacy}_i\,(t_1))^2\,[\delta_4 + \delta_5 \log y_i\,(t_1) + \delta_6 (\log y_i\,(t_1))^2] \end{aligned} \qquad (5b)$$

In (5b), I anticipate the first bracketed expression will be positive and the second negative. This captures the initial urban concentration benefits as scale economies are exploited, which then peak and are followed by losses from further concentration. While (5b) is a complex relationship and a simpler specification (e.g., δ_5, $\delta_6 = 0$) might capture essentials, the more complex relationship is strongly significant. Finally, we can recognize that the effects of urban concentration may also vary with country scale. To do that, inside either the first or both square brackets in (5b), I add a $\log(\text{scale}_i\,(t_1))$ term, with other δ coefficients.

Results are given in Table 2. I discuss just the GMM results (where coefficients are more precisely estimated). OLS and fixed effect results are reported in columns 5-6, for the specification on which I focus most of the discussion. In the first specification in column 1, from equation (5a), the effects of primacy vary with income. In this expression, in estimation in Table 2, as expected $\delta_1 > 0$ and $\delta_2 < 0$, so as income rises initially that works to make increases in urban concentration either helpful or, at least, less harmful. In the GMM estimates in column 1 the (log) income point which maximizes any positive effect of primacy on growth ($-\delta_1/(2\delta_2)$) equals 6.8, which is a GDP per capita (1987 PPP$) of about $1000.

However at this point the whole expression in square brackets in equation (5a) is still negative. That means increases in concentration are always harmful, just less so around a GDP per capita of $1000. But this specification doesn't allow there to be too much or too little urban concentration. Column 2 captures that. In column 2, I estimate equation (5b), when there is a quadratic primacy specification, where both terms are interacted with income (in quadratics). In the results, as anticipated initial increases in primacy are good, but there is a peak, or best point. Second, that point changes with income. Results for the column 2 specification are given in Table A1 of the Appendix. They are very similar to our primary results in column 4.

For the column 4 results, I lastly account for country size effects, where we expect the optimal degree of primacy to decline as country scale increases. I estimate equation (5b) adding in a country scale term, as measured by log (national urban population). Column 3 shows results with scale interacted with both primacy and primacy squared. Since the term with primacy squared is not significant at the 5% level, in column 4, I report results when scale is only interacted with primacy. Results discussed in the text are based on column 4 coefficients and that is the specification for which I report OLS and fixed effect results. Results on best primacy points for column 3 coefficients in Table 2 are in Table A2 of the Appendix. These results, with the strong quadratic form to scale effects, produce noticeable regions of low income-scale space where we have no identifiable best concentration level. However, they may give a better sense of best primacy for larger countries. I also experimented with adding to the primacy terms the interaction between income and scale, but it is not significant.

Excessive Urban Concentration and Its Costs.

There are two sets of outcomes from the estimates in column 4 of Table 2. First, for the point estimates, in Table 3, I calculate (1) what the best primacy value is at different income and country scale levels, (2) what the *standard error* is for that point calculated applying the delta method to the best primacy point function, and (3) what the gain in annual growth rate is in moving from one standard deviation above (or, below) the best point, to the best point. In the second set of results later on, I examine which countries tend to have excessive concentration levels.

In Table 3 the best primacy points, standard errors, and growth losses of excessive urban concentration are calculated for different income values and different national urban scales (8 million urban population, 22 million and 100 million). For these calculations, I use a standard deviation of primacy of 0.11, (lower than the overall standard deviation of 0.15, but more consistent with the sub-groups being identified (by income and below by population and income)). Note that income is purchasing power parity, so countries like India and Bangladesh both have incomes in excess of $1400 in 1995. Examples of countries at each income level in the 1990's are given in the table.

Table 3 indicates that the best primacy value does increase with income up to about $4900 and then it declines. But, for a wide interval ($1800 - 8100^+$), the best value changes little. The loss from excessive primacy is distinctly lower up to about $3000 and then peaks in the middle income ranges of $5000 - $10000, before declining modestly. In Table 3, these growth losses at a given income level don't vary with country scale. From the calculations based on Table 2 column 3 coefficients, Table A2 results in the Appendix suggest they do vary with country scale, declining modestly as country scale rises.

A key result is that the income losses, at any income level, of excessive concentration are substantial. The point estimates suggest losses in annual (percentage) growth rates of income of up to 1.5. With annual percentage growth rates over the time period across countries averaging a little over two, these are very large losses. Alternatively viewed, they are of similar magnitude to the effect on growth of a one standard deviation increase in human capital or in the investment rate. They suggest that the concerns in the literature concerning the national resources potentially squandered due to excessive urban concentration are valid.

In Figure 1, I graph this primacy, growth, and income relationship, showing the marginal effect on 5-year growth rates of different income-primacy pairs, for a small-medium size country. The size is 13400 urban residents, the anti-log of the mean log(scale) (which is considerably less than the mean of scale itself). Part A shows the three dimensional relationship, indicating the quadratic effect of primacy on growth and its quadratic relationship to income. The latter implies, although it is visually hard to read, that the best primacy point varies with income. In part B, to clarify, I graph the growth-primacy quadratic relationship, showing how the best primacy point first shifts out as income rises and then modestly shifts back.

In terms of other results in Table 3, for any income level, as expected, the best concentration level falls with country size. Second, the error bands on best primacy values once income rises above $1400 are quite tight, which will allow us to better next identify countries with significantly too little or too much primacy.

To further illustrate the point, in Part B of Table 3 for hypothetical low, medium, and high income countries, I examine growth losses from excessive primacy for a medium-size country (22 million urban residents) from Table 2, column 4 estimates. I assign average realized growth rates over 1960-1995 to the group with excessive primacy (best primacy levels plus 1 (within group) standard deviation of primacy: 0.11). Note this value of excessive primacy falls outside the 95% error bands on best primacy values for the reported size-income pairs. For best primacy countries I add in the differential growth, attributable to improved primacy. Growth losses both absolutely (in annual growth rate points) and relatively (between poor and best primacy countries) rise with income levels. For a low income country over or under concentration reduces growth rates by 0.9 from 5.8, while in a high income country the reduction is 1.6 from 3.6.

The quantitative, but not qualitative results in Tables 2 and 3 can vary somewhat with choice of covariates. An obvious question, given the quadratic form to income we use to interact with primacy variables, concerns whether a quadratic income term alone is itself significant. Such a variable has a zero coefficient, with no influence on any other results. If I replace national urban population with national population as the scale measure, results on best primacy points are very similar to those in Table 2. However the region of income-scale pairs where increases in primacy are always harmful enlarges considerably. Similarly, the fixed effect results in column 6 of Table 2 relative to those in columns 3 or 4 suggest larger regions of parameter space where primacy is always harmful and even lower best primacy points. If I add in Barro's inflation and government consumption variables or I add in the degree of openness to trade, the sign patterns and magnitudes of primacy and all its interaction effects are unchanged. I don't rely on a Barro specification because the variable list is so long that the GMM procedure in DPD98 can only utilize instrument lists for one period (versus at least two prior periods of values), reducing the efficiency of estimation. Finally, as footnoted earlier, substantially altering the instrument list has no pronounced effect on results.

Other general comments on the results in Table 2 are that with primacy under GMM, speeds of convergence are even higher than in Table 1. For average income and primacy, annual speeds of convergence are 10-15%! Most other variables maintain similar results to those in Table 1, with fertility and schooling displaying some sensitivity to specification. I also note that variables considered in the analysis of primacy below such as the degree of federalism, or political decentralization, and the density of transport infrastructure have no significant direct effects on growth.

Who Has Excessive Urban Concentration?

Which and how many countries operate with substantially excessive urban concentration? I have 79 countries and examine their situation in 1990. In 1990, seven of these countries have missing data, so we looking at 72 countries. I calculate from column 4 in Table 2 the best primacy level for each country, given country income and urban population in 1990. Then, I compare it with actual 1990 primacy. In Table 4, I define a country as having too much or too little primacy if it is more than two standard deviations above or below the best primacy point for that country. If it is within two standard deviations, I say it has satisfactory primacy. I then list the countries which fall into each category, **if** they meet an additional criterion which is that they do not contradict results from Table 2, column 3 (see the footnote in the table). Also, countries in Table 4 in parentheses are those for which Table 2 column 3 results suggest increases in primacy are always harmful. In naming the names, I want to restrict consideration to countries that meet multiple criteria for the category to which they are assigned!

There are 30 countries with satisfactory urban concentration. Examples of countries with satisfactory urban concentration in 1990, include federal countries such as USA, Canada, Australia, and New Zealand. Six of the countries with satisfactory primacy appear to have more excessive primacy (at least .08 above the best primacy point) under column 3 results (Ghana, Honduras, Indonesia, Mexico, Philippines, and Zimbabwe). There are 24 countries under the basic column 4 results with excessive urban concentration, which is 34% of the countries with best identifiable primacy points in 1990. The list with highly excessive primacy includes the usual suspects – Latin American countries such as Argentina, Panama, Costa Rica, and Chile, Asian countries such as Korea and Thailand, African countries such as Congo, and European countries like Greece, Portugal and Ireland. Many are countries with more explicitly unitary governments, or where federal structures have been severely constrained traditionally.

Finally, based on Table 2 column 4 coefficients, there are also a substantial number of countries--16-- with too little primacy and two where there is no identified best primacy level. However Table 2 column 3 results contradict this finding of too little primacy in eleven cases, suggesting primacy is satisfactory (or, even potentially excessive--Brazil) in those cases (Brazil, Bulgaria, W. Germany, Italy, Poland, South Africa, Spain, Sweden, Switzerland, U.K., Venezuela). So I list examples of five countries with too little primacy. These include Belgium (a small, split country) and special cases such as the former Yugoslavia and Czechoslovakia or such as Malaysia with Singapore carved off.

2. **Determinants of Urban Concentration**

What are the determinants of urban concentration? Urban concentration is not a growth process converging to some steady-state value. Rather it is expected to differ across time and countries with country size and level of economic development. In particular, given the growth results on "best primacy" and the literature which suggests those results, I expect urban concentration to initially increase as income rises, to peak, and then to decline with further increases in income. I also expect urban concentration to decline with country scale – national urban population. One would also expect urban concentration might be lower in countries with large land areas, where resources are spread out. Or it may be higher if the primate city is a port, benefiting from international trade.

Apart from these natural market, scale, and geographic features, urban concentration will be influenced by policy/institutional variables. For example, urban concentration would be expected to decrease as the degree of federalism in a country increases. Federalism tends to level the playing field for competition across cities. Hinterland states and cities have more autonomy to provide their own services and infrastructure investments so as to attract firms and workers from primate cities. Similarly, if the primate city is a national capital, that may increase urban concentration, given tendencies of national governments to favor their national capitals with special services and infrastructure investments.

Alternatively, or in addition to institutional measures such as federalism, we can look at measures of interregional transport infrastructure investments that open up coastal markets to hinterland producers. We have measures of the density of national navigable waterways (which don't really change over the time period) and measures of the density of the national road system. Such investments are expected to lower urban concentration.

Finally, as a reflection of policies, there is the degree of imports plus exports in GDP. The expected impact of increased openness on urban concentration is ambiguous. On the one hand, given primate cities may be coastal ports and/or centers of international commerce, increases in openness may favor the primate city. On the other hand, following the new economic geography literature (Fujita, et al. (1999)), increases in openness may open up international markets to hinterland producers and allow them to compete more effectively with primate cities. We will see what the empirical evidence suggests.

Specification and Data

Urban concentration in country i in time t is specified as

$$\text{Primacy}_i\,(t) = a + X_i\,(t)\,B + f_i + \eta_t + \varepsilon_{it} \qquad (10)$$

The $X_i(t)$ are covariates suggested in the above discussion, f_i a country fixed effect, η_t a time dummy and ε_{it} a contemporaneous error term. In final specifications, I use a mix of lagged covariates $X_i\,(t\text{-}1)$, and contemporaneous ones, as explained below. In estimating (10), I start with OLS. But under OLS the $X_i\,(t)$ are not likely to be exogenous to the fixed f_i, or even the ε_{it}. Accordingly, following the growth econometric implementation, I first difference (10) to get

$$\text{Primacy}_i\,(t) - \text{Primacy}_i\,(t\text{-}1) = (X_i\,(t) - X_i\,(t\text{-}1))\,B + (\eta_t - \eta_{t\text{-}1}) + (\varepsilon_{it}\ \varepsilon_{it\text{-}1}) \qquad (11)$$

Equation (11) is estimated by GMM, where each differenced time period is a separate equation – year, with the β coefficients constrained to be equal across time periods. $X_i\,(t\text{-}2)$, $X_i\,(t\text{-}3)$ and so on backwards in time are used as instruments (along with national land area). Estimated equations pass Sargan test with flying colors, but there is some evidence of second degree serial correlation implying level error terms (e.g., ε_{it} and $\varepsilon_{it\text{-}1}$) are correlated. I presume any such correlation is moving average serial correlation, but there could be a problem that the underlying stochastic process may be more complicated. Also, in general, the results on the primacy equation estimation are much less robust, than those for the growth equation; and this section should be read with that in mind.

To recover the effects of time invariant variables in (10) from GMM estimations, I regress the residuals (Primacy$_i$ (t) – X$_i$ (t) $\hat{B}$) on a constant term, time dummies, and the variables. I treat national land area, kilometers of navigable waterways, and whether the primate city is a port or national capital as time invariant and exogenous. Enough countries have each permutation of national capital or not and port or not to identify separate port and capital effects. In the sample estimating period, no country changes whether the primate city is a port or not and only two countries have the primate city change whether it is a national capital or not. The assumption that these time-invariant variables are exogenous is of course a stretch. A variety of unmeasured time invariant country characteristics may influence choice of capital or historical determination of waterways. So the results from the residuals regressions are only suggestive.

In terms of data, for the additional variables to the previous section, openness ((exports plus imports)/GNP) is from the Penn World Tables. For transport, I have (time invariant) kilometers of navigable waterways from the CIA World Factbook and time varying kilometers of roads (motorways, autobahns, highways, and main national, secondary and regional roads) from the International Road Federation supplemented by CIA data, for 1967 and 1970-1995. Both measures are divided by national land area. Given I control for national urban population and given per person road investments are much higher in rural areas, I am presuming the variation in national road densities should capture investments in interregional road systems. I also measured transport infrastructure by highway density, but the definition of this variable is much less consistent across countries. Results are quite similar and I footnote them.

Finally, a federalism variable was constructed for all countries over 10 million population in 1990, with the variable having values for 1960-1995. The index increases from zero to four representing increasing degrees of local autonomy. The index is based on nine categories. For the first four, countries get a value of either zero or four if (1) the government structure is officially unitary or federal (2) the regional executive is not elected or elected (3) the municipal government is not elected or elected and (4) the national government can suspend local or regional governments or not. Categories 5 [and 6] take values zero, two, or four if provincial and local governments have no revenue raising [expenditure] authority, limited authority, or more complete authority. Categories 7, 8, and 9 for primary education, infrastructure provision and policy give values from 0-4 depending on whether provision is entirely central, mixes of central, regional, and local, or all local. These values are averaged for each country across the nine [or fewer where relevant] categories. A second index was constructed on just the first six categories. I report results for the second index, since information on categories 7-9 is noisy. But results for the two are very similar.

Results

I present the urban concentration results in Table 5 in subsections: (a) results with economic-geography variables (b) results for institutional/political variables and (c) results on policy variables such as openness and transport infrastructure. All specifications control for time dummies, but in GMM estimation those tend to be of inconsistent sign over time and generally insignificant. There is no evidence of changing world trends towards increasing or decreasing urban concentration. Second, in deciding whether to use contemporaneous or lagged covariates, in experiments contemporaneous values of income and national scale (the non-policy variables) always dominate lagged values. However for policy and institutional variables, this is not the case and lagged values offer more consistent results. The obvious reason may be that institutional and policy changes affect primacy with a lag. In results, for economic-geography variables I use X(t)'s and for institutional-policy ones, X(t-1)'s.

In Table 5, I focus on the GMM results in columns 1-4, with OLS results reported in column 5 for the GMM specification in column 4. Fixed effect results for this last specification do not have significant coefficients for the basic covariates and are not reported. Column 1 is the simplest (and robust) specification, with income, scale and openness determining primacy, along with the time "invariant" variables national land area, density of national waterways, and port and capital status.   Coefficients for these latter variables in GMM are determined from residuals regressions and are in italics. Columns 2 and 3 experiment with the federalism and road density variables. Column 4 turns to my preferred specification, with fairly robust interactive effects.   These allow for openness effects to interact with port status and, most critically, for road effects to interact with income.

Economic-Geography

In Table 5 urban concentration declines with country urban scale, although the magnitudes under GMM can be small. But for the column 4 results, with a coefficient of -.027, a one-standard deviation (2) increase in country scale decreases primacy by .05, about

1/3 standard deviation (.15) of primacy, a noticeable effect. Primacy increases as income rises from low income levels but then peaks and decreases. The peak varies but in column 4 it is about $2400, almost the same as that in Wheaton and Shishido (1981). This supports the Williamson hypothesis, from a positive perspective.

For geographic variables, national land area reduces urban concentration, as a country's resources are spread spatially. Under GMM, the residuals regressions suggest a large effect, .06, or about 40% of a standard deviation of primacy, again a noticeable effect. Until column 4, primacy increases by about 33% of a standard deviation if the primate city is a port – the benefits of being a port and "international city". In column 4 when port is interacted with openness of the economy, as discussed below, the raw effect of port is more than halved.

Political-Institutional Variables

If the capital city is the primate city that enhances primacy as anticipated, by about 40% of one-standard deviation of primacy in all specifications. This may suggest that national capitals are the beneficiary of special treatment. Federalism, as anticipated reduces urban concentration, but the effect is surprisingly small. Moreover, in the OLS version of column (2) (not reported) which exploits cross-country (not just time) variation in the data, federalism effects are even weaker! In column 2, a one-standard deviation (1) increase in the index reduces primacy by about 4% of a standard deviation of primacy. Since only 40 countries in estimating samples have federalism data (countries over 10m population in 1990), I don't include the variable in other specifications. When it is included in a specification with, say, road density, while still significant, its coefficient is halved. Since the direct effects of federalism seem very small, it might appear that federalism plays a more indirect role, through determination of whether national capitals are primate cities or the extent of historical investment in inter-regional transport infrastructure, and through federalism's relationship to national land area. However in OLS and GMM, exclusion of other such variables only raises the effect of a one standard deviation increase in the degree of federalism to, at most, a reduction in primacy equal to 7% of a standard deviation of primacy. Recall also that federalism has no direct effect on economic growth rates. Finally I note (cf. Ades and Glaeser (1995)) that Freedom House indices of the degree of democratization have no consistent effect on primacy.

Policy Variables

Countries which invested historically in navigable waterways have reduced urban concentration. A one-standard deviation (.018) increase in waterway density reduces primacy by about 20% of a standard deviation of primacy. Similarly higher road density appears to reduce primacy. A one-standard deviation increase in roads (1.5) reduces primacy by about 10% of one-standard deviation in column 3. The effect of openness in columns 1-3 is generally negative – the economic geography hypothesis – but not always (e.g., column 1). It became apparent that road infrastructure, openness, geography and income interacted in more complex fashions and that to tease out the effect of openness and infrastructure investments required more thought. Column 4 is the result of a variety of experiments.

First, it seems that the effect of openness should depend on whether the primate city is a port and, hence, already tilted towards international markets. Otherwise, increased trade is more likely to have the hypothesized effect from the economic geography literature-- to help hinterlands by opening up international markets to them. In column 4, a one-standard deviation (37) increase in openness **de**creases primacy by 8% of a standard deviation of primacy. But if the primate city is a port, a one-standard deviation increase in openness **in**creases primacy modestly (4% of a primacy standard deviation). These results critically control for road density. Perhaps the key result is that the effects of changes in openness on primacy are fairly modest, especially when the primate city is a port. Attempts to interact openness with road density or income produced non-robust results.

The most compelling new result in examining interactions turned out again to be related, in some sense, to the Williamson hypothesis. While investment in road infrastructure reduces urban concentration, the magnitude of the effect depends on income. Also accounting for the income interaction is critical to assessing any magnitudes of the effects of inter-regional investment in roads on primacy. For example, from the column 4 coefficients, a one standard deviation increase in road density (1.5) reduces primacy at incomes of $850, 4900, 13,400, and 17,500 by respectively 13, 18, 19 and 18 percent of a standard deviation of primacy. The peak is at $13,500, where the reductions are significantly larger than at low income levels.[5] How is this related to Williamson? At low income levels when there are

[5] Using highway density instead of road density, the coefficients corresponding to those in column 4 for transport density and then that interacted with income and income squared are (all significant) .411, -.103 and .0063. Here the greatest impact of highway density in reducing density is at an income of $3500 where a one

forces promoting spatial concentration in the economy, the effect of increasing road density is lower than at higher income levels when there are forces promoting dispersion. As a policy consideration, this takes heightened importance once we consider the fact that the growth losses from excessive urban concentration tend to rise with income. I expand on this notion now, conducting a simple policy experiment that ties together the growth and primacy results.

A Policy Experiment.

What is the indirect effect then on national growth rates of investment in roads through its effect on urban concentration, at different income levels? In growth regressions, after accounting for national investment rates and primacy effects, there is no direct effect of road investments on economic growth. I pick income levels of $1100, 4900, and 13400 for a medium-size, national urban population. I start primacy at a high level-- one standard deviation above its best value for the given income and scale, as in Table 3, Part B. I then ask what is the effect on annual growth rates of a one standard deviation increase in national road density. Such investment will reduce urban concentration towards its best level.

Working through the calculations in Tables 2 and 5, the results are in Table 6. At the lower income level, the effect of expanded roads in reducing primacy is to add .23 of a percentage point to the national annual economic growth rate. But at higher income levels, the effect is tripled, adding .68 of a percentage point to the annual national growth rate. That is a very large indirect effect of road expansion. Of the variables we have considered including the degree of federalism and openness of the economy, the clearest policy instrument in reducing urban concentration is expansion of interregional transport systems opening up hinterland markets. But I haven't represented the cost side of such a policy.

3. **Conclusions**

This paper argues that, if urban over-concentration is really the problem much of the literature suggests that it is, it ought to show up as affecting economic growth rates. This paper explores this issue econometrically, using a panel of 80-100 countries every 5 years from 1960 to 1995. There are three main sets of findings.

At any level of development there is indeed a best degree of national urban concentration. The best degree increases sharply as income rises up to a per capita income of

---

standard deviation increase in density reduces primacy by about 8% of a standard deviation of primacy-- a about $5000 (Penn World Table purchasing parity income), before declining modestly. The best degree of concentration declines with country scale. Growth losses from significantly non-optimal concentration are large. These losses tend to rise with level of development, peaking at a very high level of about 1.5 annual percent points of economic growth. Results are very robust.

In a group of 72 countries in 1990, approximately 30 have satisfactory concentration, 24 have noticeably excessive concentration, and 5-16 countries have too little. The list with **highly** excessive includes the usual suspects – Latin American countries such as Argentina, Panama, Costa Rica, and Chile, Asian countries such as Korea and Thailand, African countries such as Congo, and European countries like Greece, Portugal and Ireland. Many are countries with more explicitly unitary governments, or where federal structures have been severely constrained traditionally. Countries with too little urban concentration include Belgium (a small, split country) and special cases such as the former Yugoslavia and Czechoslovakia.

In terms of the determinants of urban concentration, it declines with national scale; and, in a positive version of Williamson, it initially rises with income and then peaks around an income of $3000, before declining. If the primate city is a port, increased trade leads to increased urban concentration; otherwise increased trade leads to deconcentration as hinterland markets open up with trade. However trade effects are modest. Similarly, the effects of a greater degree of political decentralization, or increased federalism, while significantly reducing urban concentration, are quite modest.

The key policy variable here affecting concentration is investment in inter-regional transport infrastructure. In particular, increases in the density of road networks significantly reduce concentration, with the effect rising with income. That is, at low income levels when there are Williamson forces promoting spatial concentration in the economy, the effect of increasing road density is lower than at higher income levels, when there are forces promoting dispersion. This fact assumes heightened importance because the growth losses from excessive concentration tend to rise with income. To illustrate, I consider the effect on national growth rates of investment in roads, through its effect on concentration. I start concentration at one standard deviation above its best value and ask what is the effect on annual growth rates of a one standard deviation increase in national road density. At the

---

smaller effect than with road density.

higher income levels, the effect of expanded roads in reducing concentration is to add .68 of a percentage point to the growth rate, about three times the effect at low income levels.

References

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

Arellano, M. and S.R. Bond (1991), "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations," Review of Economic Studies, 58, 277-297.

Barro, R. (1991), "Economic Growth in a Cross-Section of Countries," Quarterly Journal of Economics, 106, 407-443.

Barro, R. (1997), Determinants of Economic Growth, MIT Press, Cambridge.

Becker, R. and J.V. Henderson (1999), "Intra-industry Specialization and Urban Development" in Economics of Cities, Huriot and J.-F. Thisse (eds.), Cambridge University Press.

Black, D. and J.V. Henderson (1999), "A Theory of Urban Growth," Journal of Political Economy, 107, 252-284.

Caselli, F., E. Esquire, and F. Lefort (1996), "Re-opening the Convergence Debate: A New Look at Cross-Country Growth Empirics," Journal of Economic Growth, 1, 363-389.

Cass, D. (1965), "Optimal Growth in an Aggregate Model of Capital Accumulation," Review of Economic Studies, 32, 233-240.

Ciccone, P. and R. Hall (1995), "Productivity and Density of Economic Activity," American Economic Review, 86, 54-70.

Durlauf, S. and D. Quah (1998), "The New Empirics of Economic Growth," NBER Working Paper No. 6422.

Fujita, M., P. Krugman, and A. Venables (1999), The Spatial Economy, Cambridge: MIT Press.

Gabiax, X. (1999), "Zipf's Law for Cities: An Explanation," Quarterly Journal of Economics.

Glaeser, E., H.D. Kallal, J. Scheinkman and A. Shleifer (1992), "Growth in Cities," Journal of Political Economy, 100, 1126-52.

Hansen, N. (1990), "Impacts of Small and Intermediate-Sized Cities on Population Distribution: Issues and Responses," Regional Development Dialogue, Spring, 11, 60-76.

Henderson, J.V. (1974), "The Sizes and Types of Cities," American Economic Review, 64, 640-56.

Henderson, J.V. (1988), Urban Development: Theory, Fact and Illusion, Oxford University Press.

Henderson, J.V. (1999a), "Marshall's Scale Economies," NBER Working Paper #7358.

Henderson, J.V. (1999b), "Notes on the Costs of Urban Primacy" Brown University mimeo (10-24-99).

Henderson, J.V. and A. Kuncoro (1996), "Industrial Centralization in Indonesia," World Bank Economic Review, 10, 513-540.

Islam, N. (1995), "Growth Empirics: A Panel Approach," Quarterly Journal of Economics, 110, 1127-1170.

Kaiser, K. (1999), "Pre- and Post-Liberalization Manufacturing Location in Indonesia (1975-96)," 5-27-99 mimeo, LSE.

Kolko, J. (1999), "Can I Get Some Service Here: Transport Costs, Cities, and the Geography of Service Industries," Harvard University mimeo. (March).

Lucas, R.E., Jr. (1988), "On the Mechanics of Economic Development," Journal of Monetary Economics, 22, 3-42.

Mankiw, G., D. Romer and D. Weil (1992), "A Contribution to the Empirics of Economic Growth," Quarterly Journal of Economics, 107, 407-437.

Renaud, B. (1981), National Urbanization Policy in Developing Countries, Oxford University Press.

Richardson, H. (1987), “The Costs of Urbanization: A Four-Country Comparison,” Economic Development and Cultural Change, 33, 561-580.

Tolley, G., J. Gardner and P. Graves (1979), Urban Growth Policy in a Market Economy, New York: Academic Press.

UN (1993), World Urbanization Prospects: The 1992 Revision, New York: United Nations.

Wheaton, W. and H. Shishido (1981), "Urban Concentration, Agglomeration Economies, and the Level of Economic Development," Economic Development and Cultural Change, 30, 17-30.

Williamson, J. (1965), “Regional Inequality and the Process of National Development,” Economic Development and Cultural Change, June, 3-45.

World Bank (2000), Entering the 21st Century World Development Report 1999/2000, Oxford University Press.

Table 1. Basic Growth Results
(dependent variable: $\log y_i(t_2) - \log y_i(t_1)$)
(standard errors in parentheses)

	OLS		Fixed effects		GMM	
	(1)	(2)	(1)	(2)	(1)	(2)
log (y_i (t))	-.066**	-.087**	-.298**	-.348**	-.248**	-.399**
	(.011)	(.012)	(.036)	(.034)	(.020)	(.0089)
annual avg. rate of investment (share of GDP)	.0055**	.0055**	.0087**	.0071**	.0072**	.0070**
	(.00084)	(.00096)	(.0014)	(.0016)	(.00093)	(.00056)
log (average fertility rate)	-.092**	-.101**	-.185**	-.215**	-.046	-.115**
	(.020)	(.021)	(.046)	(.094)	(.034)	(.019)
average years of high school and college of the population over 25	.012	.015*	.018	.015	.071**	.053**
	(.0089)	(.0088)	(.017)	(.019)	(.013)	(.012)
annual avg. inflation rate		-.000059**		-.000062**		-.000082**
		(.000026)		(.000030)		(.000005)
annual avg. rate of government consumption (share of GDP)		-.0021**		-.0054**		-.010**
		(.00085)		(.0018)		(.00058)

N [countries]	679	577	679 [106]	577 [98]	467 [103]	382 [92]
adj R^2	.258	.277				
Sargan test [df, p-value]					71.81 [66, .291]	84.04 [94, .760]
annual speed of convergence	1.4%	1.8%	7.1%	8.6%	5.7%	10.2%

** significant at 5% level
* significant at 10% level

Table 2. Growth and Urban Concentration

	GMM				OLS	fixed effect
	(1)	(2)	(3)	(4)	(5)	(6)
Primacy*						
(δ_0	-5.93**	-42.8**	-41.4**	-43.3**	-12.2**	-19.4**
	(2.18)	(4.77)	(7.83)	(6.55)	(4.77)	(7.96)
δ_1 log y_{t1}	1.72**	10.2**	11.1**	11.0**	2.86**	5.28**
	(.522)	(1.18)	(2.12)	(1.78)	(1.17)	(2.06)
δ_2 (log $y_{t1})^2$	-.013**	-.569**	-.606**	-.615**	-.172**	-.331**
	(.031)	(.078)	(.136)	(.124)	(.071)	(.134)
δ_3 log scale)			-.683**	-.363**	.061**	-.170**
			(.252)	(.048)	(.019)	(.081)
Primacy sq.*						
(δ_4		70.2**	54.5**	63.3**	13.9	33.3*
		(8.90)	(14.9)	(14.3)	(9.94)	(18.5)
δ_5 log y_{t1}		-16.4**	-13.5**	-15.1**	-3.20	-8.72*
		(2.46)	(4.23)	(3.99)	(2.41)	(4.88)
δ_6 (log $y_{t1})^2$		.889**	.713**	.828**	.181	.555*
		(.169)	(.292)	(.280)	(.145)	(.324)
δ_7 log scale)			.401			
			(.400)			

	(1)	(2)	(3)	(4)	(5)	(6)
log $y_i(t_1)$	-.314**	-.586**	-.775**	-.724**	-.095**	-.347**
	(.039)	(.052)	(.111)	(.084)	(.026)	(.083)
invest rate	.0084**	.0088**	.011**	.011**	.0050**	.0089**
	(.0012)	(.0015)	(.0016)	(.0016)	(.0010)	(.0015)
log (fertility	.049	-.068	-.165**	-.106**	-.088**	-.184**
rate	(.038)	(.050)	(.034)	(.031)	(.027)	(.049)
avg. years	.103**	.113**	.041**	.073**	.025**	.0020
hs college	(.016)	(.014)	(.020)	(.015)	(.0082)	(.021)
R^2					.331	.364
N [countries]	359 [79]	359 [79]	359 [79]	359 [79]	518 [80]	518 [80]
Sargan	73.9	65.2	63.6	67.5		
test	[87, .840]	[88, .967]	[106, 1.0]	[97, .990]		
[d.f., p-value]						

In column 4, if standard errors from step one of the GMM procedure are used, all variables remain significant at the 5% level except the last two in the primacy sq. expression which have standard errors of 9.2 and .61 respectively.

Table 3. Urban Concentration Effects as Income and Country Size Change

Best Primacy Values

[Standard errors for best primacy, *annual growth rate losses from excessive primacy*]

A. Table 2, Column 4 Results

National Urban Population	8 million	22 million	100 million
Income: GDP per capita (purchasing power parity: Penn Tables)			
850 (Mali, Mozambique, Sudan)	.15 [.078, *.47*]	.042 [.148, *.47*]	neg
1100 (Cameroon, Senegal, Jordan)	.22 [.045, *.71*]	.15 [.072, *.71*]	.034 [.096, *.71*]
1800 (Egypt, Indonesia, El Salvador)	.26 [.023, *1.1*]	.22 [.033, *1.1*]	.15 [.033, *1.1*]
3000 (Colombia, Thailand, Ecuador)	.28 [.017, *1.4*]	.24 [.022, *1.4*]	.18 [.022, *1.4*]
4900 (Malaysia, Hungary, Argentina)	.28 [.016, *1.6*]	.25 [.018, *1.6*]	.20 [.023, *1.6*]
8100 (Portugal, Greece, Spain)	.28 [.016, *1.6*]	.24 [.018, *1.6*]	.20 [.018, *1.6*]
13400 (Belgium, W. Germany, Japan)	.27 [.019, *1.6*]	.23 [.021, *1.6*]	.18 [.021, *1.6*]
17200 (USA)	.26 [.020, *1.5*]	.22 [.027, *1.5*]	.17 [.042, *1.5*]

B. Annual Growth Rates and Urban Concentration *

(medium size country)

	Best Primacy	Excessive Primacy (+ lsd (.11))
Low Income ($1400)	5.8%	4.9
Median Income ($4900)	5.8	3.9
High Income ($13,400)	3.6	2.0

*The calculation gives excessive primacy countries the average growth rate among different income groups (low: $GDP_{pc} < \$3000$; high: $GDP_{pc} > 8100$). It then augments that growth rate by the added growth from being at best primacy (rather than excessive primacy).

Table 4. Countries with Excessive
Or Too Little Concentration in 1990*

No. of countries with Satisfactory primacy, (with examples)	30 (Algeria, Australia, Bolivia, Cameroon, Canada, (China), Colombia, Denmark, Ecuador, Finland, Hungary, (India), Iran, Jordan, (Kenya), (Mozambique), New Zealand, (Pakistan), (Sudan), Syria, Tunisia, Turkey, (Zambia), U.S.A.)
No. of countries with Excessive primacy	24 (Argentina, Austria, Bangladesh, Chile, Congo, Costa Rica, Dominican Republic, Egypt, El Salvador, France, Greece, Guatemala, Ireland, Israel, Japan, Korea, Nicaragua, Panama, Paraguay, Peru, Portugal, Senegal, Thailand, Uruguay)
No. of countries with Too little primacy	16 (Belgium, Czechoslovakia, Malaysia, Netherlands, Yugoslavia)
No. of countries With no best Identifiable Primacy point	2 (Mali, Uganda)

---

* A country is not by name listed if based on Table 2 column 3 results it doesn't fall in the assigned category. For that I required satisfactory countries to remain within .08 of best primacy and excessive or under concentration countries to remain .08 beyond their best value. Countries in parentheses have no best primacy value under Table 2, column 3 results.

Table 5. Determinants of Urban Concentration

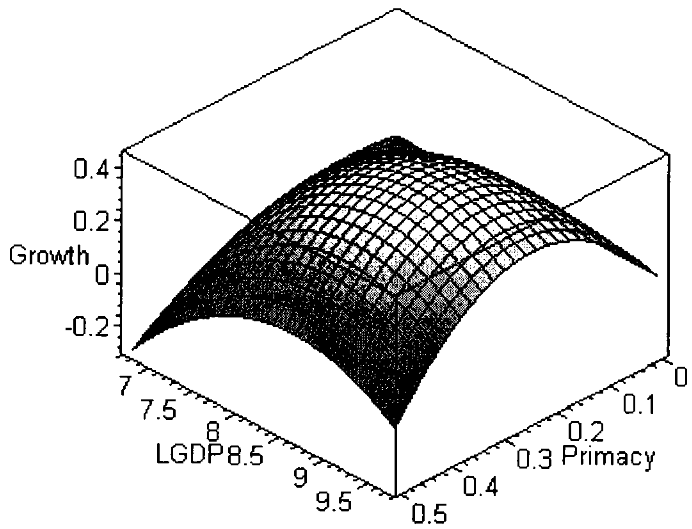
	GMM				OLS
	(1)	(2)	(3)	(4)	(5)
log (nat. urban pop., t)	-.0057** (.0028)	-.056** (.0072)	-.0069** (.0016)	-.026** (.0041)	-.052** (.0063)
log (GDP per cap., t)	.217** (.017)	.163** (.043)	.094** (.016)	.101** (.033)	.219 (.143)
log (GDP pc, t) ** 2	-.017** (.0012)	.0091** (.0028)	-.0055** (.00098)	-.0065** (.0020)	-.013 (.0087)
openness, t-1	.00071** (.00025)	-.00018** (.000032)	-.000085** (.000011)	-.00032** (.000038)	-.0014** (.00022)
port * openness, t-1				.00050** (.000044)	.0017** (.00033)
port city	*.070*** *(.0094)*	*.045*** *(.013)*	*.060*** *(.011)*	*.030*** *(.010)*	-.036* (.021)
capital city	*.101*** *(.010)*	*.085*** *(.011)*	*.110*** *(.0097)*	*.089*** *(.0091)*	.076** (.010)
log (nat. land area)	*-.032*** *(.0028)*	*-.0037*** *(.0042)*	*-.035*** *(.0034)*	*-.027*** *(.0032)*	-.015** (.0067)
log (nat. water way density)	*-1.83*** *(.159)*	*-1.09*** *(.176)*	*-1.92*** *(.221)*	*-1.57*** *(.169)*	-1.32** (.184)
federalism index, t-1		-.0061** (.0013)			

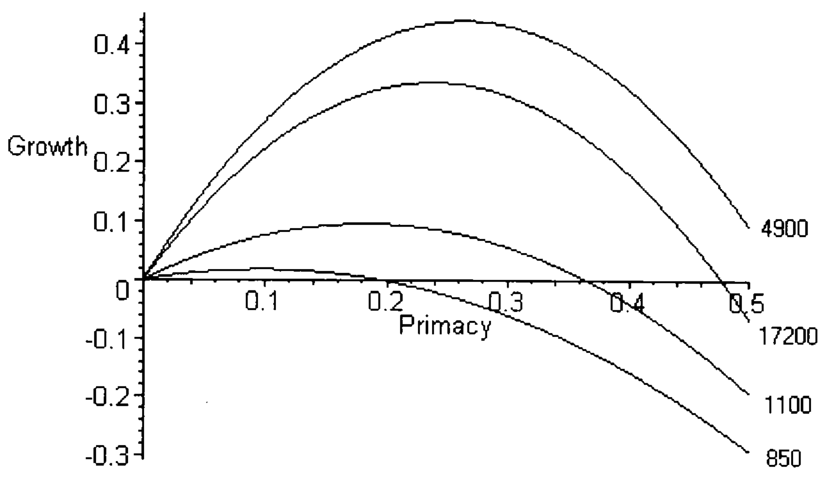
	(1)	(2)	(3)	(4)	(5)
log (road density, t-1)			-.013** (.0016)	.072** (.036)	.173 (.157)
log (road dens., t-1) * log (GDP pc, t)				-.020** (.0082)	-.049 (.041)
log (road dens., t-1) * log (GDP pc, t)**2				.0011** (.00046)	.0032 (.0026)
time dummies	yes	yes	yes	yes	yes
Adj. R^2					.594
Sargan value [d.f., p-value]	83.1 [86, .568]	38.9 [106, 1.0]	74.7 [75, .487]	67.9 [117, 1.0]	
N [countries]	510 [92]	247 [40]	316 [77]	316 [77]	415 [88]

Table 6. Increase in Annual Economic Growth Rate of Increased National Road Density.
(One standard deviation increase in road density in a medium size country with excessive primacy)

Low income ($1100)	.23 percentage point
Medium income (4900)	.68
High income (13400)	.68

Figure 1: Urban Concentration and Growth





Appendix

Table A1. Urban Concentration Effects as Income Changes
(No country size effects)

Income: GDP per capita (purchasing power parity: Penn Tables)	best primacy value	loss in annual growth rate from 1 s.d. excessive primacy (in percentage points)
$850	0.58	0.71%
1100	.232	.33
1800	.255	.77
3000	.256	1.1
4900	.251	1.3
8100	.242	1.5
13400	.227	1.5
17200	.217	1.4

Table A2. Urban Concentration Effects as Income and Country Size Change

Best Primacy Values [and Growth Losses from Excessive Primacy]

Table 2, Column 3 Results

Income	National Population: 8million	22 million	100 million
850	neg	neg	neg
1100	.07 [.38]	neg	neg
1800	.21 [.78]	.12 [.68]	neg
3000	.25 [1.1]	.19 [1.0]	.071 [.83]
4900	.26 [1.3]	.21 [1.2]	.12 [1.1]
8100	.25 [1.5]	.21 [1.3]	.13 [1.2]
13400	.24 [1.5]	.20 [1.4]	.12 [1.2]
17200	.23 [1.5]	.18 [1.4]	.10 [1.2]

Table A3. Descriptive Statistics

	Mean	Standard deviation
Log (GDP per capita)	7.97	.962
GDPpc growth rate (5 yr.)	.110	.150
Avg. annual investment rate	18.2	8.96
Log (avg. fertility rate)	1.34	.527
Avg. years of high school and college, pop. over 25	1.17	1.06
Avg. annual inflation rate	61.4	387
Avg. an. govt. cons. rate	17.5	7.48
Primacy	.305	.154
Log (nat. urban pop.)	8.77	1.45
Openness	62.7	36.8
Log (nat. land area)	12.1	2.12
Log (nat. waterway density)	.0079	.018
Federalism index	1.32	1.27
Log (nat. road density)	-1.75	1.46