The Output Cost of Latin America's Infrastructure Gap

César Calderón
Central Bank of Chile

Luis Servén
The World Bank

This Version: July 2002

Abstract

Latin American has lost substantial ground relative to other developing and developed regions in terms of the quality and quantity of infrastructure over the last two decades. For instance, Latin America's infrastructure gap in infrastructure growth relative to that of the seven successful economies of East Asia grew by 40-50 percent for road length, 50-60 percent for telecommunications, and as much as 90-100 percent in terms of power generation capacity over the 1980-97 period.

The consequences of this loss of ground for growth and welfare in the region are a matter of concern. Lack of adequate infrastructure services results in lower productivity and higher production costs for private producers. Poor road and telecommunication networks raise transport and, more generally, logistic costs, which have been shown in comparative studies to exceed the international norm by wide margins (Guasch, 2001). The reduced profitability in turn discourages private investment. Through all these channels, the result is lower output growth. As with infrastructure, Latin America's loss of ground was particularly marked in the 1980s.

We find evidence of a strong empirical association between output and infrastructure although this need not reflect causation from infrastructure services to aggregate output. The present paper wants to determine the role of Latin America's growing infrastructure gap in the widening of the output gap. We devote our efforts in this paper to answer this question.

JEL Classification: E23, O54
Key Words: Infrastructure, Output Gap, Latin America

1. Introduction

Over the last two decades Latin America lost substantial ground vis-à-vis other developing and developed regions in terms of the quality and quantity of infrastructure assets. While there was considerable diversity across countries in the magnitude of this phenomenon, it affected virtually all infrastructure sectors in all of the region's countries.[1]

Table 1 provides a summary illustration of Latin America's infrastructure growth relative to that of the seven successful economies of East Asia.[2] The table presents the change in the infrastructure gap – measured by East Asia's infrastructure stocks per worker relative to those of Latin America – over 1980-97, using both regional averages and medians.

[Insert Table 1 here]

The two sets of figures tell the same story: Latin America's infrastructure gap grew by a huge margin in the last two decades: 40 to 50 percent for road length, 50 to 60 percent for telecommunications (defined as the total number of main telephone lines), and as much as 90-100 percent in terms of power generation capacity. The loss of ground was particularly marked in the 1980s in terms of all three assets in the table. In the 1990s, Latin America continued to fall behind at a rapid pace in power generation capacity, but its loss of ground in terms of transport routes proceeded at a slower pace than in the previous decade, and the gap in terms of telecommunications infrastructure ceased to expand.[3]

The consequences of this loss of ground for growth and welfare in the region are a matter of concern. Lack of adequate infrastructure services results in lower productivity and higher production costs for private producers. Poor road and telecommunication networks raise transport and, more generally, logistic costs, which have been shown in comparative studies to exceed the international norm by wide margins (Guasch 2001). The reduced profitability in turn discourages private investment. Through all these channels, the result is lower output growth. For later reference, the bottom of Table 1 also shows that the gap in GDP per worker (in PPP-adjusted terms) between East Asia and Latin America grew by some 90 percent over 1980-97. As with infrastructure, Latin America's loss of ground was particularly marked in the 1980s.

Figure 1 brings out graphically the association between infrastructure accumulation and growth performance. The figure plots the average growth rate of GDP per worker over the last four decades against the average rate of growth of infrastructure endowments – with the latter measured by the simple average of the growth rates of phone lines, kilometers of roads and power generation capacity, all in per worker terms. Even with this crude measure of infrastructure stocks, a strong positive correlation between

---

[1] See Calderón, Easterly and Servén (2002).

[2] Hong-Kong, Indonesia, Korea, Malaysia, Taiwan, Thailand and Singapore.

[3] However, if we look at total (main + mobile) phone lines rather than just main lines, the relative performance of Latin America in the 1990s would be somewhat less favorable than shown in the table.

infrastructure accumulation and growth performance is apparent. Indeed, a simple cross-country regression of growth on infrastructure accumulation yields a highly significant positive regression coefficient and an R-squared of 32 percent.

[Insert Figure 3.1 here]

Of course, strong as this empirical association is, it need not reflect causation from infrastructure services to aggregate output. In fact, the observed correlation could actually reflect reverse causation from GDP to infrastructure demand, or the action of third factors affecting both GDP and infrastructure stocks. Thus, the key question is: what was the role of Latin America's growing infrastructure gap in the widening of the output gap? The rest of this paper is devoted to answering that question.

2. Methodological Approach

Our empirical approach is based on the estimation of an aggregate production function augmented with infrastructure capital. The analysis is closely related to that in Canning (1999), and follows a recent literature concerned with the contribution of infrastructure to aggregate output (Canning and Bennathan 2000, Demetriades and Mamuneas 2000, and Röller and Waverman 2001).

For simplicity, we adopt a Cobb-Douglas specification of the infrastructure-augmented production function:[4]

$$y = \alpha k + \beta h + \gamma z + (1 - \alpha - \beta - \gamma) l \qquad (1)$$

where y is aggregate value added (GDP), k is the physical non-infrastructure capital stock, l denotes labor, h is human capital and z is a measure of infrastructure capital. All the variables are expressed in logs, and constant returns to scale are assumed.

Importantly, (1) implicitly assumes that infrastructure services are a fixed proportion of the infrastructure capital stock. Thus, other things equal, larger stocks should be reflected in higher aggregate output. This approach is analogous to that conventionally used in standard production functions excluding infrastructure, which assume that physical and human capital services are proportional to the respective stocks k and h.

In principle, the parameter γ in (1) should capture the elasticity of output with respect to infrastructure for given values of the other inputs. However, this presumes that k includes non-infrastructure capital only. In reality, what we have is data on the total capital stock, including both infrastructure and other physical assets. Thus, infrastructure capital appears twice in the equations –as part of k, and separately as z. Hence, the parameter γ captures the extent to which the productivity of infrastructure exceeds (if $\gamma > 0$) or falls

---

[4] Canning and Benathan (2000) and Demetriades and Mamuneas (2000) also present estimates using translog specifications.

short of (γ <0) the productivity of non-infrastructure capital; see Canning (1999) for further discussion.

The contribution of infrastructure capital to output can be found by noting that the measured capital stock is a weighted sum of infrastructure and other physical assets, with weights given by their respective relative prices. Thus, letting $\tilde{k}$ denote non-infrastructure physical capital, we can write:

$$k \approx \frac{\tilde{K}}{\tilde{K} + p_z Z} \tilde{k} + \frac{p_z Z}{\tilde{K} + p_z Z} z \qquad (2)$$

where uppercase letters denote the anti-logs of lowercase variables; p_z is the relative price of infrastructure capital in terms of non-infrastructure capital, and we have assumed that the latter is approximately equal to the price of overall capital, under the presumption that infrastructure assets are typically a small fraction of the total capital stock.[5]

Combining (1) and (2), the elasticity of output with respect to infrastructure can be expressed:

$$\frac{\partial y}{\partial z} = g + qa \equiv h_z \qquad (3)$$

where

$$q \equiv \frac{p_z Z}{\tilde{K} + p_z Z} \qquad (4)$$

is the share of infrastructure in the overall physical capital stock. These expressions involve log-linear approximations around an arbitrary point (e.g., the sample mean), and hence q should be evaluated accordingly. In practice, since infrastructure stocks typically account for relatively small portions of the overall capital stock, the difference between h_z and the 'naïve' estimate g should be fairly modest.

Finally, it is worth noting that (4) only captures the direct impact of infrastructure on output, leaving aside the possible indirect impact occurring through the effects of infrastructure on the accumulation of other productive inputs – most importantly, non-infrastructure capital. To the extent that both types of capital are gross complements in production (as assumed here), an increase in infrastructure capital raises the profitability of non-infrastructure capital and, other things equal, should lead over time to a higher $\tilde{K}$, which in turn should cause a further output expansion. By ignoring this indirect effect, we are likely underestimating the contribution of infrastructure to output over the long term.[6]

---

[5] A similar procedure is followed by Canning and Bennathan (2000).

[6] On this see Demetriades and Mamuneas (2000), who distinguish between the 'short run', with non-infrastructure capital predetermined, and the 'long run', over which non-infrastructure capital adjusts to its optimal value. They also define an 'intermediate run' in which the capital stock partially adjusts to its equilibrium level.

3. Empirical Implementation

For estimation purposes, equation (1) above is rewritten in terms of ratios to the labor force:

$$y_{it} - l_{it} = a_i + b_t + a(k_{it} - l_{it}) + b(h_{it} - l_{it}) + g(z_{it} - l_{it}) + e_{it} \qquad (5)$$

Here the subscripts i and t are used to index countries and years, respectively; the terms a_i, b_t capture country-specific and time-specific productivity factors, and e_{it} is a random disturbance that will be assumed uncorrelated across countries and over time.

Our objective is to estimate the parameters of equation (5) using a large panel data set. We use annual data for the period 1960-97 from 101 industrial and developing countries– close to 4,000 observations in total. In practice, some of the instrumental variable estimators employed below use up several lags of the variables to construct instruments, so that we end up with 101 countries and 3,232 observations. To ensure comparability across estimators, we limit ourselves to this reduced sample even when employing simpler estimators using no lags.[7]

Sample coverage and data sources are described in detail in the Appendix. In essence, the measures of output (GDP) and physical capital per worker that we use are based on suitably expanded versions of the Summer-Heston dataset, while the (log) human capital stock is measured by the number of years of secondary schooling of the working-age population.[8]

Regarding infrastructure capital, we focus primarily on the three standard indicators of infrastructure endowments used in Table 1: (i) electricity generating capacity (in GW), (ii) road length (in Km.), and (iii) the number of main telephone lines. However, we also perform some experiments with alternative measures of infrastructure capital. We scale each of these variables dividing by the total labor force. While these measures of infrastructure capital are admittedly crude –in particular, they do not capture variations in the quality of infrastructure –we choose them in view of their broad availability across countries and over time, and their frequent use in the recent empirical growth literature.

There is by now a considerable literature reporting empirical estimates of equations similar to (5) above (see Gramlich 1994 for an overview). In the present panel context, there are four main issues to take into consideration: cross-country heterogeneity, common factors, measurement error and endogeneity.

---

[7] However, empirical estimates using the entire sample are very similar to those using the reduced sample.

[8] This choice is consistent with the findings in Barro and Sala-I-Martín (1995). Empirical specifications using broader definitions of human capital (inclusive of primary and/or tertiary schooling) yield more imprecise estimates of the contribution of human capital, and have only minimal effects on the coefficients of the physical capital and infrastructure variables. To save space, they are not reported.

The first issue relates to possible cross-country heterogeneity of the production technology. Imposing a common technology when in reality production functions vary across countries would lead to inconsistent estimates. To address this issue, we allow for country-specific effects a_i in the estimations below. Indeed, omission of fixed effects is known to lead to a large overstatement of the contribution of infrastructure to output (see e.g. Holtz-Eakin 1994; Roller and Waverman 2001).

A second specification issue concerns the possible existence of omitted common factors – such as the world business cycle – causing output to move together across countries. These common factors can result in cross-country residual correlation, which in turn would lead to invalid inferences with the estimation methods that we shall be using below. To eliminate the common factors, we allow for time-specific effects in the estimated regressions; this is equivalent to a regression in which each variable enters as deviation from its cross-sectional mean in the year in question.

The third problem is measurement error, which is likely to be important particularly in the case of infrastructure stocks. There are two reasons for this. On the one hand, the quality of the stocks (e.g., the condition and capacity of roads; the reliability of power plants, and so on) can vary greatly not only across countries but also within countries. Unfortunately, data on the quality of infrastructure is not readily available for a large cross-country time-series dataset such as the one under consideration.[9] On the other hand, the timing of changes to the reported infrastructure stocks is to some extent arbitrary – e.g., impassable roads or unusable portions of railway track may remain in the books for some time, before being suddenly removed from the reported stock figures; new power plants may not become fully operational until some time after completion, and so on. Formally, these considerations imply that infrastructure may be measured with error, so that the time-varying disturbance e_{it} may include a measurement error correlated with the infrastructure variables. Standard estimates of (5) would therefore be subject to attrition bias, most likely causing underestimation of the coefficients of the infrastructure stocks.

Related to this is the problem of endogeneity that may affect the infrastructure regressors in (3.5) – and perhaps also the physical and human capital stocks per worker. It can be argued that infrastructure stocks are jointly determined with output per worker – in fact, the positive correlation of infrastructure stocks with output found in the data could just reflect the fact that the income elasticity of infrastructure demand is positive. Arguably, similar considerations could be made for the physical and human capital stocks.

In the univariate case, standard least-squares estimation in the presence of reverse causation from output to infrastructure would lead to an upward bias in the infrastructure coefficient; in the multivariate case the situation is more complex and the direction of the bias cannot be established a priori – and even more so in the presence of measurement error that may introduce attenuation bias.

---

[9] Note that by including time and country effects in the empirical specification we can account for country-specific levels, as well as cross-country changes, in infrastructure quality – but not for country-specific quality changes.

One way to address the two-way causality between infrastructure and output would be to develop a fully-specified simultaneous model of infrastructure supply and demand. Unfortunately, this would pose stringent data requirements well beyond the scope of this research.[10]

An alternative, less demanding way to tackle both measurement error and endogeneity is to use an instrumental-variable estimation approach. However, there are few exogenous instruments available with the broad time-series and cross-country coverage that we need here. Demographic variables are perhaps the only obvious source of identifying information – since they are likely to affect the demand for infrastructure (as well as physical and human capital) services without being subject to reverse causation or correlated with the infrastructure measurement error. Thus, we use urban population and population density (both in logs) as outside instruments.

We complement these strictly exogenous instruments with appropriate ‘internal’ (i.e., weakly exogenous) instruments constructed along the lines of Griliches and Hausman (1986) and Arellano and Bond (1991), given by suitably lagged values of the explanatory variables in (5). Specifically, we take first differences of (5) to remove the country-specific effect:[11]

$$\Delta\left(y_{it}-l_{it}\right)=c_t+a\Delta(k_{it}-l_{it})+b\Delta(h_{it}-l_{it})+g\Delta(z_{it}-l_{it})+\Delta e_{it} \tag{6}$$

where $c_t=b_t-b_{t-1}$. Under appropriate assumptions about the serial correlation of e_{it} (the time-varying disturbance, possibly inclusive of measurement error), lagged levels of the right-hand side variables become valid instruments. In particular, if e_{it} is serially uncorrelated and the regressors are weakly exogenous (that is, uncorrelated with future realizations of e_{it}, but not with its current or past realizations) then the second and higher lags of the regressors become valid instruments in (6). More generally, if e_{it} follows a moving average process of order q, then lags q+2 and higher of the regressors would become valid instruments.

Validity of the instruments used in the estimation can be tested directly through Sargan tests of orthogonality between the instruments and the error term, as well as indirectly through tests of first- and higher-order autocorrelation of the errors, see Arellano and Bond (1991). For example, if e_{it} is serially uncorrelated, then its first difference included in (6) should display first but no higher-order autocorrelation, in which case twice-lagged regressors are indeed valid instruments, as stated earlier.

---

[10] In particular, one would need cross-country time-series data on the prices of infrastructure services, which are not available for a broad country sample such as the one considered here. The only example of such an approach in the recent literature is Roller and Waverman (2001), who develop an empirical supply-demand model along the lines in the text but including only telecommunications infrastructure. The model is estimated using data for OECD economies.

[11] Note that lagged levels of the right-hand side variables are unlikely to provide valid instruments for the estimation of (3.5) due to the presence of time-invariant country-specific factors which may be correlated with the levels of the regressors at all lags.

The above discussion characterizes the difference-GMM estimator of Arellano and Bond (1991). However, under additional assumptions[12] a more efficient IV estimator may be available, namely the system GMM estimator of Blundell and Bond (1998), which combines estimation of (6) and (5) using lagged differences of the regressors as instruments in the level equation (5). The validity of these additional instruments can be checked through difference-Sargan tests of orthogonality between the extra instruments and the error term.

4. Estimation results

Table 2 reports the sample correlations among the dependent and independent variables. The figures below the main diagonal reflect the correlation among the levels of the variables, while those above the diagonal correspond to their first differences. Anticipating some of the experiments below, we present two alternative infrastructure measures for transport routes, total roads and total roads plus railways (with the latter variable available only for a smaller country sample[13]), and two measures as well for telecommunications -- main lines and total lines, including cellular.

In both levels and differences, real GDP per worker shows a significant correlation with each of the infrastructure measures, as well as with the physical and human capital stocks per worker. Among the infrastructure variables, the biggest correlation with GDP corresponds by far to the telecommunication measures. Unsurprisingly, the magnitude of the correlations is much bigger when the variables are expressed in levels than when they are expressed in differences. In turn, the various infrastructure measures are also positively correlated with each other, again more so in terms of levels than in terms of differences. Finally, there seems to be little difference between the two alternative measures of transport routes (their correlation exceeds 0.99 in both levels and differences) and the two measures of telecommunications infrastructure (their correlation is 0.97 in differences and virtually 1.00 in levels).

[Insert Table 2]

Before proceeding to GMM estimation, Table 3 reports empirical results using simpler estimators for equation (5).[14] The first two columns present OLS estimates on the cross-section (column 1) and pooled sample (column 2), neither of which is robust to heterogeneity, measurement error or endogeneity of the regressors. The two sets of estimates are in fact quite similar: in both cases we find a sizable output contribution of the capital stock, and a significant effect of telecommunications infrastructure. The

---

[12] For lagged differences of a regressor x to provide a valid instrument for the levels equation, we need $E[a_i x_{it}] = E[a_i x_{is}]$ for all *t* and *s*. This is essentially a stationarity assumption; see Blundell and Bond (1998).

[13] Railway data are unavailable for some 300 country-year observations.

[14] Except for column 1, all estimates reported in this and later tables include a full set of year dummies that resulted highly significant in all cases.

remaining coefficients are insignificant, although that on transport routes approaches statistical significance with a counter-intuitive negative sign. The pooled OLS results also show strong evidence of serial correlation of the residuals, a clear symptom of misspecification.

Column 3 reports the within estimator, which controls for country-specific effects but not for endogeneity or measurement error. In the presence of the latter, the within transformation can lead to quite misleading estimates, see Griliches and Hausman 1986. In the present case, it can be seen that all the regressors carry positive coefficients, all significantly different from zero except for that of transport routes. Among the infrastructure variables, telecommunications carries a much larger coefficient than the rest, similarly to the OLS results.[15]

The estimators presented so far ignore the issues of measurement error and endogeneity. Column 4 reports 2SLS estimates of (5) using as instruments the current and first three lags of urban population and population density, plus the second lags of the explanatory variables.[16] The estimates obtained in this manner are quite similar to the pooled OLS estimates, and equally disappointing. Apart from the physical capital stock, only the telecommunications variable is significant. Moreover, a Sargan test rejects the validity of the instrument with a p-value of less than .001 – an unsurprising outcome in view of the strong evidence of autocorrelation of the residuals shown in the table, which provides a clear indication of misspecification.

[Insert Table 3 here]

Table 4 turns to GMM estimation using alternative specifications and instrument sets. Column 1 reports our base specification, using the difference GMM estimator and the same instrument set as in the last column of Table 3 -- twice-lagged levels of the explanatory variables plus the current value and three lags of the exogenous demographic variables. Comparison of these GMM estimates with the within estimates in Table 3.3 shows that in every case the former are larger in magnitude than the latter – which hints at the possible presence of attenuation bias in the within estimates. Moreover, the GMM estimates of the coefficients of all three infrastructure variables are all statistically significant (although only at the 10 percent level in the case of power). They are also of roughly similar magnitude. Finally, the diagnostic tests provide support for the selected specification – the Sargan test shows no evidence against the validity of the instruments and, as anticipated, the serial correlation tests hint at first-order but no higher-order serial correlation of the differenced error term.

---

[15] We also computed various panel cointegration estimates, using the techniques of Kao and Chiang (2000) for nonstationary panels, with results very similar to the within estimates in table 3 (see also Baltagi 2000). These estimates are subject to the same measurement error and simultaneity biases as the within estimator. They are not reported here to save space.

[16] In anticipation of other experiments reported later, the instrument set includes also second lags of primary and tertiary schooling, total roads per worker, and total phone lines per worker.

Column 2 provides a robustness check by lagging the instruments one extra period – that is, using the third rather than the second lags of the regressors as instruments (in addition to the demographic variables). The results are virtually identical to those in the preceding column, and the diagnostic tests continue to lend support to the specification.

So far we have been using lagged infrastructure and physical capital stocks as instruments. One might object that these variables could belong in the production function, so that they do not provide identifying information. We can test this assertion by dropping them and limiting the instrument set to the exogenous demographic variables. Thus in column 3 of Table 4 we include as instruments only use the current and first two lags of urban population and population density, as well as their squares, along with the second lag of the schooling variables. Nevertheless, the estimation results are fairly similar to those in the preceding columns. The only exception is the coefficient on power generating capacity, which becomes considerably larger than before. All other coefficients are virtually unchanged, although that on roads is now estimated with poor precision.

Finally, in column 4 we turn to the system GMM estimator of Blundell and Bond (1998), combining the levels equation (3.5) with the first-difference equation (3.6), and adding as instruments for the former the twice-lagged first differences of the same instruments used in column 1. The resulting parameter estimates are fairly different from those obtained from the difference-GMM estimator. If anything, they are close to the within estimates in the previous table. However, the Sargan test clearly rejects the validity of the instruments, while the difference-Sargan test (not shown in the table) yields a p-value of less than .001 percent and thus provides an equally strong indication of misspecification. This suggests that the stationarity condition discussed earlier, required for the validity of the system GMM estimator, does not hold in our data.

[Insert Table 4 here]

In view of these results, we base our remaining experiments on the difference-GMM estimator and retain the same set of instruments as in the base specification of column 1 in Table 4. Using this as a starting point, in Table 5 we experiment with alternative specifications. Column 1 just reproduces the initial specification for ease of comparison. In column 2, we use roads plus railways, rather than roads alone, to summarize the transport network infrastructure. This leads to the loss of some 10 percent of the sample. The parameter estimate on the combined transport variable is quite similar to that obtained earlier using roads only, although the point estimate is somewhat imprecise. As for the other parameters, the coefficient on power rises about 50 percent relative to its value in column 1, while that on phone lines declines somewhat. However, these changes are modest relative to the standard errors. The other coefficients remain unchanged.

Next, in column 3 we replace main phone lines with total (main + mobile) phone lines as our indicator of telecommunications infrastructure. This makes virtually no difference for any of the parameter estimates, nor for the diagnostic statistics, all of which are almost identical to those in column 1.

Finally, in column 4 we look for nonlinear effects of telecommunications equipment, along the lines reported in Roller and Waverman (2001), who find that the elasticity of output to telecommunications stocks rises with the telecommunications stock. To explore this issue, we add in the equation the square of main phone lines per worker. Its estimated coefficient turns out to be negative, but wholly insignificant, while the remaining coefficients show virtually no change. Hence, we conclude that our data show little indication of nonlinear effects of telecommunications infrastructure.[17]

In all the specifications reported in Table 3.5, the diagnostic statistics are quite supportive of the model. The Sargan tests show no evidence against the choice of instruments, and the serial correlation tests provide a mild suggestion of first- but no higher-order autocorrelation.

[Insert Table 3.5 here]

5. The output cost of Latin America's infrastructure gap

As noted earlier, the empirical estimates reported so far do not capture the total contribution of infrastructure to output, because infrastructure stocks are already included in the overall capital stock. To identify that impact, we need to compute the elasticity of output with respect to infrastructure assets as in equations (3)-(4).

To compute the share of the different infrastructure stocks in the overall capital stock, we use data on the cost of infrastructure assets collected by Canning and Bennathan (2000). There are some caveats, however. These costs are available only for a limited number of countries, and do not necessarily correspond to assets of homogeneous quality. They also show a great degree of cross-country variation. For our purposes, since we are primarily interested in the performance of Latin America, we compute the capital stock shares using the cost data available for countries in this region and the average ratios of the relevant stocks over 1980-97; we then take the medians of the country-specific figures. Limited experiments with alternative ways to construct these shares usually led to roughly similar results; however, since many other procedures are possible, the results have to be taken as illustrative. They are reported in the middle column of Table 6.

According to the figures in the table, telecommunications infrastructure accounts for just over 1 percent of the overall capital stock, while power and roads represent 14 and 16 percent, respectively. Using these shares for the calculation in equation (4), we obtain the elasticities reported in the third column of the table. As it turns out, the elasticities of the three infrastructure stocks are all of similar magnitude, with the largest corresponding to

---

[17] It is also useful to compare these estimates with the results of Roller and Waverman (2001) for OECD countries. Their production function specification ignores human capital, roads and power, does not impose constant returns, and employs a nonlinear transformation of the stock of phone lines. It can be shown that if the same transformation were used here, then the resulting estimate of the elasticity of output with respect to phone lines would be very similar to that reported by Roller and Waverman. The elasticity with respect to physical capital, however, is much higher in their case (over .50).

roads and the smallest to phone lines. The differences are very small, however – on the order of a few hundredths of a percent -- and in view of the uncertainties surrounding the underlying calculations, we opt for using a common value below for all three, which as a working hypothesis we place at 0.16.

[Insert Table 6 here]

This estimated elasticity can be used to provide a rough idea of the contribution of infrastructure stocks to the diverging performance of GDP per worker between Latin America and the East Asian tigers over the last two decades. More precisely, we calculate the portion of the change in the gap in GDP per worker between the two regions that can be attributed to the differential evolution of their respective infrastructure stocks -- the infrastructure gap – that was portrayed in Table 1 above.

This is done in Table 7, which shows the role of each infrastructure asset in the widening GDP gap, as well as the combined role of all three vis-à-vis the other inputs – human capital and non-infrastructure physical capital. The table reports calculations using both unweighted means and regional medians.

The estimated contributions of the infrastructure assets are substantial. The top line in the table shows that all three assets combined account for about one-third of the widening GDP gap between East Asia and Latin America. In other words, the differential evolution of infrastructure assets in Latin America and East Asia widened the cross-regional gap in GDP per worker by some 30 percent over 1980-97.

Of this total, the largest contribution (nearly half) corresponds to power generating capacity, while phone lines and roads combined had an impact of similar magnitude to that of power infrastructure on the GDP gap. This relative ranking of assets is unsurprising in view of their respective evolution depicted earlier in Table 1, according to which power had the worst performance over the two decades under analysis. It is worth noting also that the results are very similar whether regional medians or averages are employed in the calculation.

The table also shows the contributions of the two conventional inputs – physical (non-infrastructure) and human capital. The slower accumulation of physical capital in Latin America relative to East Asia accounts for another 30 percent increase in the output gap – an amount similar to that attributable to infrastructure. Finally, the differential evolution of human capital across the two regions is responsible for up to another 10 percent increase in the output gap.

Finally, the bottom line in the table shows that the estimated model tends to under-predict the change in the output gap between the two regions. Between 15 and 20 percent of the latter is left unexplained.

[Insert Table 7 here]

Table 8 offers an individual-country perspective on the same phenomenon. For each country, the table reports the change in the infrastructure gap and the income gap (vis-à-vis the East Asia average) over 1980-97, as well as the contribution of the former to the latter. The first three columns of the table show that over the period in question nearly every country in Latin America lost ground relative to East Asia in all three infrastructure assets considered. The only exceptions were Chile and Jamaica in telecom, and Uruguay in roads. On the other hand, every country listed in the table lost ground in terms of power generation capacity per head, and the extent of the lag was particularly dramatic in Panama, Guatemala, Nicaragua and the Dominican Republic. Except for Panama, these countries were also the least dynamic in terms of the stock of roads, while Panama, Ecuador and Mexico lost the most ground in telecom.

The table also shows that the contribution of the infrastructure gap to the gap in income per worker – computed in the same way as in the preceding table –was positive for every country listed. In other words, in every country the widening infrastructure gap added to the income gap over the sample period. The output cost of lagging infrastructure was particularly large in Central America: in Nicaragua, Guatemala and Panama the loss of ground in terms of infrastructure assets widened the income gap by over 40 percent relative to East Asia. At the other end, Jamaica and Uruguay were the least bad performers – i.e., their loss of ground in infrastructure only involved a relatively modest cost in terms of output per worker.

[Insert Table 8 here]

6. Summary and Conclusions

Over the last twenty years Latin America fell behind in terms of infrastructure quantity and quality vis-à-vis other developing and industrial regions. Virtually all countries and infrastructure sectors in the region were affected by this relative slowdown, which was particularly pronounced in the 1980s.

The analysis in this paper shows that this widening infrastructure gap can account for a considerable fraction of the increase in Latin America’s output gap relative the successful East Asian economies over the 1980s and 1990s. Lagging telecommunication assets, power generation capacity and road networks all contributed to Latin America’s loss of ground in terms of output per worker. While there is a fair degree of diversity across the region’s economies regarding the magnitude of this effect, in every one of the countries analyzed we find that lagging infrastructure added to the output lag vis-à-vis the East Asian tigers.

These conclusions are based on empirical estimates of the contribution of infrastructure stocks to aggregate output computed over a large cross-country time-series data set using an infrastructure-augmented production function specification. In this framework, we find positive and significant output contributions of all three infrastructure assets considered – as well as physical and human capital.

This approach poses some difficulties, however, such as the potential endogeneity of infrastructure stocks and the fact that they are subject to measurement error – due among other things to heterogeneity in infrastructure quality across countries and over time. We have attempted to overcome these problems using instrumental variable estimators combining internal and external instruments. On the whole, the empirical results are supportive of this approach. We find little evidence against the validity of the instruments, and the estimates do not change significantly when alternative instrument sets are used or the instrument set is restricted to exogenous demographic variables only. We take this as evidence that the empirical estimates capture the effect of the exogenous component of infrastructure on output.

Appendix
Sample Coverage and Data

In order to estimate the production functions presented in Tables 2-5, we collected annual data for 101 countries over the 1960-97 period (38 observations per country). Note that in our regression framework all figures are expressed as magnitudes per worker.

Output has been approximated by using the real GDP in 1990 PPP US dollars from Summers-Heston (1991) and complemented by the data on the Global Development Network Growth Database created by William Easterly at the World Bank. Analogously, we used data on domestic capital stock from Summers-Heston and Easterly. The labor input is proxied by the total labor force as reported by the World Bank's World Development Indicators (WDI).

Regarding infrastructure stocks, we use physical indicators for the different infrastructure sectors. First, we use the number of telephone main lines as a proxy for infrastructure in Telecommunications. We complemented the data in Canning (1998) with recent figures from the International Telecommunications Union (ITU) annual reports. Second, our proxy for infrastructure in power is the data on electricity generating capacity (in kilowatts). The main source for these data is the United Nations' Energy Statistics and Statistical Yearbook. Finally, we use data on road length (in km.) for the transportation sector. We obtained the data from the International Road Federation's (IRF) World Road Statistics. One caveat regarding these data, as noted by Canning (1999), is that they may exhibit significant variations in quality. In particular, they do not reflect the width of the roads nor their condition.

References

Arellano, M. and S. 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.

Baltagi, B. (ed.) (2000): Nonstationary panels, panel cointegration and dynamic panels. Amsterdam: Elsevier Science.

Barro, R. and X. Sala-I-Martín (1995): Economic Growth. McGraw-Hill.

Blundell, R. and S. Bond (1998):”Initial conditions and moment restrictions in dynamic panel data models”, Journal of Econometrics.

Calderón, C., W. Easterly and L. Servén (2002): “How did Latin America’s infrastructure fare in the era of macroeconomic crises?” The World Bank, Mimeo

Canning, D. (1999): “Infrastructure’s contribution to aggregate output”, World Bank Policy Research Discussion Paper 2246.

Canning, D. and E. Bennathan (2000): “The social rate of return on infrastructure investment”, World Bank Policy Research Discussion Paper 2390.

Demetriades, P. and T. Mamuneas (2000): “Intertemporal output and employment effects of public infrastructure capital: evidence from 12 OECD economies”, Economic Journal 110, 687-712.

Gramlich, E. (1994): “Infrastructure investment: a review essay”, Journal of Economic Literature 32, 1176-96.

Griliches, Z. and J. Hausman (1986): “Errors in variables in panel data”, Journal of Econometrics 31, 93-118.

Guasch, J. and J. Kogan (2001): “Inventories in developing countries: levels and determinants, a red flag on competitiveness and growth”, World Bank Policy Research Discussion Paper 2552.

Holtz-Eakin, D. (1994): “Infrastructure in a structural model of economic growth”, NBER Working Paper 4824.

Kao, Chihwa (1999), “Spurious Regression and Residual-Based Tests for Cointegration in Panel Data”, Journal of Econometrics, 90, 1-44

Kao, C. and M. Chiang (2000): “On the estimation and inference of a cointegrated regression in panel data”, mimeo.

Pesaran, M. Hashem, Yongcheol Shin and Ron P. Smith (1999). “Pooled Mean Group Estimation of Dynamic Heterogeneous Panels.” Journal of the American Statistical Association, 94: 621-634.

Pritchett, Lant (1996): “Where has all the education gone?”, World Bank Policy Research Discussion Paper 1581.

Röller, L. and L. Waverman (2001): “Telecommunications infrastructure and economic development: a simultaneous approach”, American Economic Review 91, 909-23.

Table 1

The widening infrastructure gap
Percentage change in relative infrastructure stocks per worker
(Latin America vs. East Asia)

	Medians by region			Simple averages by region		
	1980-97	1980-89	1990-97	1980-97	1980-89	1990-97
Main phone lines	63.58	45.86	-14.01	47.61	42.52	2.98
Power generating capacity	101.21	50.03	40.66	91.14	45.61	39.56
Roads	43.98	21.34	10.09	52.53	36.11	13.14
Memo item:						
Change in relative GDP per worker	88.89	52.66	26.60	90.24	55.75	26.55

Note: each cell in the table shows the percentage change in the stock of the respective infrastructure asset in East Asia minus the same change in Latin America.

Table 2
Sample correlations

	GDP	Physical capital	Secondary schooling	Electricity generating capacity	Roads	Transport routes	Main phone lines	Total phone lines
GDP	--	0.21 **	0.04 **	0.07 **	0.05 **	0.05 **	0.13 **	0.14 **
Physical capital	0.71 **	--	0.01	0.05 **	0.02	0.02	0.06 **	0.06 **
Secondary schooling	0.39 **	0.34 **	--	-0.05 **	-0.01	0.00	0.05 **	0.11 **
Electricity generating capacity	0.50 **	0.47 **	0.29 **	--	0.03	0.03	0.07 **	0.05 **
Roads	0.23 **	0.20 **	0.02	0.21 **	--	1.00 **	0.03	0.03
Transport routes	0.22 **	0.20 **	0.02	0.19 **	1.00 **	--	0.04 **	0.03
Main phone lines	0.58 **	0.55 **	0.30 **	0.60 **	0.27 **	0.25 **	--	0.97 **
Total phone lines	0.60 **	0.57 **	0.32 **	0.59 **	0.28 **	0.26 **	1.00 **	--

Note: All variables are measured per worker and (except for schooling) expressed in logs. Values below the main diagonal refer to the variables in levels; values above the diagonal refer to first differences. Two stars denote 5-percent level significance.

Table 3
Infrastructure-augmented production function: alternative estimates

(Dependent variable: log GDP per worker)

	1	2	3	4
Estimator	**Cross-section OLS**	**Pooled OLS**	**Within**	**2SLS**
Physical capital	0.472	0.387	0.245	0.414
	(5.324)	(7.685)	(7.199)	(7.644)
Secondary schooling	-0.005	0.016	0.135	0.017
	(0.123)	(0.474)	(2.758)	(0.492)
Electricity generating capacity	0.030	0.051	0.068	0.047
	(0.512)	(1.137)	(2.294)	(1.002)
Roads	-0.055	-0.046	0.026	-0.049
	(1.702)	(1.473)	(0.707)	(1.586)
Main phone lines	0.147	0.185	0.133	0.169
	(2.433)	(4.432)	(4.544)	(3.883)
R-squared	0.954	0.939	0.987	0.939
1st-order autocorrelation (p-value)		0.000	0.341	0.000
2nd-order autocorrelation (p-value)		0.000	0.945	0.000
Number of observations	101	3232	3232	3232
Number of countries	101	101	101	101

Note: All variables are measured per worker and (except for schooling) expressed in logs.
T-statistics in brackets are heteroskedasticity-consistent.

Table 4
Alternative GMM estimates

(Dependent variable: log GDP per worker)

	1	2	3	4
Model specification	**Differences**	**Differences**	**Differences**	**System**
Instruments	**Levels t-2**	**Levels t-3**	**Demographics**	**Levels +diffs**
Physical capital	0.363	0.361	0.351	0.222
	(10.832)	(11.034)	(7.903)	(7.867)
Secondary schooling	0.148	0.169	0.159	0.222
	(3.361)	(3.815)	(3.443)	(5.520)
Electricity generating capacity	0.112	0.123	0.177	0.109
	(1.809)	(2.148)	(2.468)	(2.970)
Roads	0.119	0.117	0.105	-0.005
	(2.197)	(2.'195)	(1.241)	(0.084)
Main phone lines	0.151	0.140	0.138	0.147
	(3.634)	(3.236)	(3.168)	(6.164)
Wald test of joint significance (p-value)	0.000	0.000	0.000	0.000
Sargan test (p-value)	0.319	0.312	0.141	0.002
1st-order autocorrelation (p-value)	0.111	0.106	0.143	0.555
2nd-order autocorrelation (p-value)	0.793	0.794	0.888	0.778
Number of observations	3232	3232	3232	3232
Number of countries	101	101	101	101

Note: All variables are measured per worker and (except for schooling) expressed in logs. Heteroskedasticity-consistent T-statistics in brackets.

Table 5
First-difference GMM estimates of alternative specifications

(Dependent variable: log GDP per worker)

	1	2	3	4
Physical capital	0.363	0.365	0.363	0.365
	(10.832)	(12.642)	(10.768)	(10.718)
Secondary schooling	0.148	0.119	0.139	0.153
	(3.361)	(2.780)	(3.274)	(2.792)
Electricity generating capacity	0.112	0.174	0.118	0.123
	(1.809)	(2.642)	(1.910)	(2.000)
Roads	0.119		0.117	0.119
	(2.197)		(2.180)	(2.109)
Roads + railways		0.116		
		(1.646)		
Main phone lines	0.151	0.113	0.161	0.152
	(3.634)	(2.284)	(3.507)	(2.832)
Total phone lines				
Main phone lines squared				-0.009
				(0.039)
Infrastructure				
Wald test of joint significance (p-value)	0.000	0.000	0.000	0.000
Sargan test (p-value)	0.319	0.607	0.329	0.261
1st-order autocorrelation (p-value)	0.111	0.115	0.12	0.110
2nd-order autocorrelation (p-value)	0.793	0.536	0.793	0.789
Number of observations	3232	2941	3232	3232
Number of countries	101	92	101	101

Note: All variables are measured per worker and (except for schooling) expressed in logs. Heteroskedasticity-consistent T-statistics in brackets.

Table 6

Elasticity of output per worker with respect to capital per worker

	Regression Estimate	Share of total capital stock	Total elasticity
Infrastructure Capital			
Main phone lines	0.152	0.012	0.156
Power generating capacity	0.112	0.140	0.163
Roads	0.119	0.163	0.178
Non-infrastructure capital	0.363	0.685	0.249

Note: capital stock shares are the medians of country values computed on the basis of cost data from Canning and Bennathan (2000) and asset stock data for Latin America.

Table 7

The infrastructure gap and the output gap

Contribution of various inputs to the change in relative GDP per worker
(Latin America vs. East Asia, 1980-97)

	Medians by region		Simple averages by region	
1. Infrastructure		33.40		30.61
Main phone lines	10.17		7.62	
Power generating capacity	16.19		14.58	
Roads	7.04		8.40	
2. Non-infrastructure capital		30.28		29.86
3. Human capital		10.88		7.07
Sum		74.56		67.53
Actual change in GDP per worker		88.90		90.24
Residual		14.33		22.71

Note: The contribution of each input to the change in relative output is calculated multiplying the change in the input by the respective output elasticity estimate. The elasticities used are those in Table 3.6.

Table 8

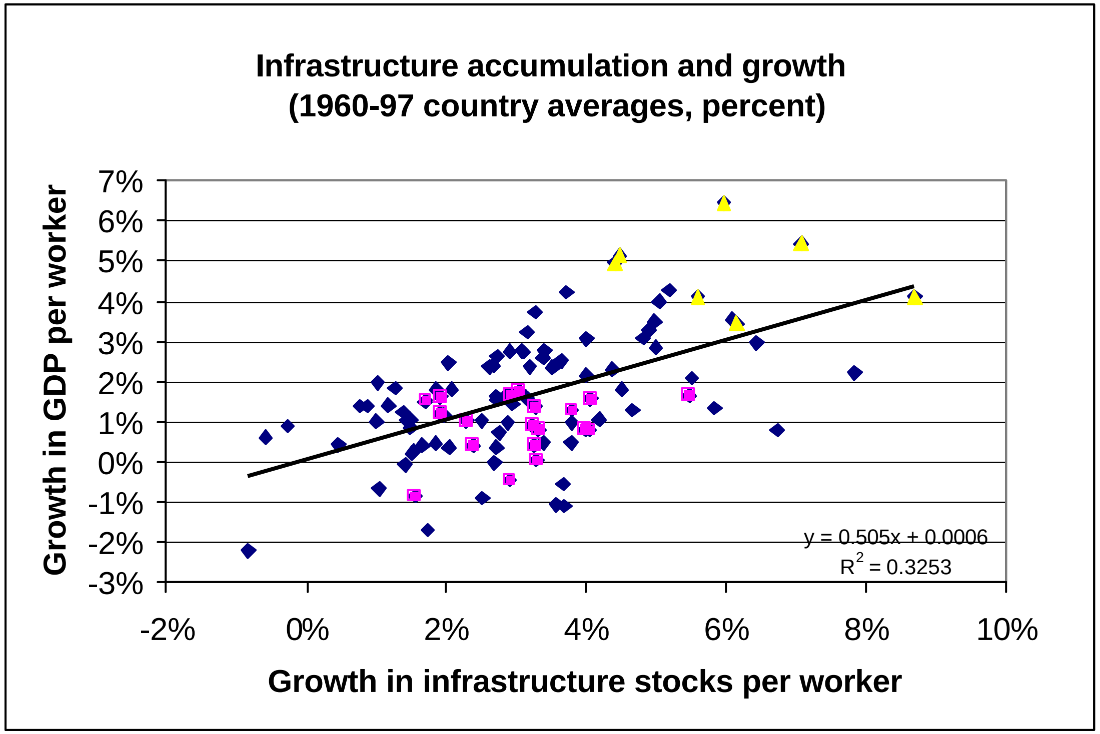
The infrastructure gap and the output gap

Contribution of the change in relative infrastructure stocks
to the change in relative GDP per worker
East Asia vs selected Latin American countries, 1980-97

Country	Relative changes per worker				Contribution of infrastructure to the change in relative output				[2]/[1]
	Infrastructure stocks			Output [1]					
	Power	Roads	Telecom		Power	Roads	Telecom	Total [2]	
Argentina	93.48%	45.84%	51.36%	86.78%	15.89%	7.79%	7.70%	31.39%	**36.17%**
Bolivia	90.29%	35.49%	56.01%	85.69%	15.35%	6.03%	8.40%	29.78%	**34.76%**
Brazil	100.75%	39.71%	71.84%	92.25%	17.13%	6.75%	10.78%	34.65%	**37.56%**
Chile	111.22%	66.66%	-5.67%	44.53%	18.91%	11.33%	-0.85%	29.39%	**66.00%**
Colombia	99.06%	48.25%	47.42%	82.46%	16.84%	8.20%	7.11%	32.16%	**39.00%**
Costa Rica	108.55%	52.76%	76.99%	87.85%	18.45%	8.97%	11.55%	38.97%	**44.36%**
Dom. Rep.	123.84%	104.15%	14.01%	82.72%	21.05%	17.71%	2.10%	40.86%	**49.39%**
Ecuador	68.45%	53.22%	80.76%	107.00%	11.64%	9.05%	12.11%	32.80%	**30.65%**
Guatemala	134.87%	97.07%	33.71%	88.85%	22.93%	16.50%	5.06%	44.49%	**50.07%**
Honduras	103.09%	65.02%	5.58%	92.22%	17.52%	11.05%	0.84%	29.42%	**31.90%**
Jamaica	100.05%	41.38%	-24.72%	98.29%	17.01%	7.04%	-3.71%	20.34%	**20.69%**
México	73.23%	33.14%	82.48%	94.15%	12.45%	5.63%	12.37%	30.45%	**32.34%**
Nicaragua	153.53%	84.19%	66.73%	113.80%	26.10%	14.31%	10.01%	50.42%	**44.31%**
Panama	140.67%	41.93%	94.40%	89.19%	23.91%	7.13%	14.16%	45.20%	**50.68%**
Peru	119.62%	48.79%	33.86%	102.92%	20.34%	8.29%	5.08%	33.71%	**32.75%**
El Salvador	90.46%	66.11%	33.83%	91.91%	15.38%	11.24%	5.07%	31.69%	**34.48%**
Uruguay	43.94%	-16.91%	49.79%	76.99%	7.47%	-2.87%	7.47%	12.06%	**15.67%**
Venezuela	97.37%	36.49%	84.21%	105.13%	16.55%	6.20%	12.63%	35.39%	**33.66%**

Note: for each country, the contribution of each infrastructure asset to the change in relative output is calculated multiplying the change in the asset stock (relative to the East Asia median) by the respective output elasticity estimate from Table 3.7.

Figure 1


Infrastructure accumulation and growth
(1960-97 country averages, percent)
Growth in GDP per worker
7%
6%
5%
4%
3%
2%
1%
0%
-1%
-2%
-3%
-2%
0%
2%
4%
6%
8%
10%
Growth in infrastructure stocks per worker
y = 0.505x + 0.0006
R2 = 0.3253