

A Spatial Explanation for the Balassa–Samuelson Effect*

Péter Karádi[†] and Miklós Koren[‡]

January 31, 2008

Abstract

We propose a model of urbanization and development to explain why the price level is higher in rich countries (the “Penn effect”). There are two sectors: manufacturing, which is freely tradable, and non-tradable services, which have to locate near customers in big cities. As countries develop, total factor productivity increases simultaneously in both sectors. However, because services compete with the urban population for scarce land, labor productivity will grow slower in services than in manufacturing. Services become more expensive, and the aggregate price level becomes higher. The model hence provides a theoretical foundation for the Balassa–Samuelson assumption that productivity growth is slower in the non-tradable sector than in the tradable sector. The key implications of the model are borne out in the data: while the Penn effect is very strong among countries where land is scarce (densely populated, highly urban countries), there is virtually no correlation between income and the price level among rural countries.

1 Introduction

Rich countries are more expensive than poor ones. This is the well-known “Penn effect.”¹ Figure *** plots the price level of the consumption basket relative to the U.S. against (PPP-adjusted) GDP per capita for *** countries in the *** benchmark year of the Penn World Tables.² Recently, Hsieh and Klenow have shown that *** Balassa and Samuelson and productivities.

Rich cities are more expensive than poor ones. According to `bestplaces.net`, the overall cost of living is 54 percent higher in Boston, MA than in Milwaukee, WI. The two cities are of similar size (close to 600,000 people), but the per capita income in Boston (\$28,000) is about 56 percent higher than that in Milwaukee (\$18,000). Similarly to the price differences

*PRELIMINARY AND INCOMPLETE. PLEASE DO NOT CIRCULATE.

[†]New York University. E-mail: peter.karadi@nyu.edu

[‡]Princeton University. E-mail: mkoren@princeton.edu.

¹The formal statement of the Penn effect is that the cost-of-living adjusted incomes differences are smaller than the unadjusted income differences.

²***details on construction

across countries, the biggest price differences are in nontradable sectors. Healthcare is 33 percent more expensive in Boston, whereas housing is about 200 percent more expensive.³ A simple explanation for the large service price differences across cities is that rents are higher in rich cities, which sustains the high service prices.

We propose a model of development and urbanization that reconciles these two facts. There are two sectors: manufacturing, which is freely tradable, and non-tradable services, which have to locate near customers in big cities. As countries develop, total factor productivity increases simultaneously in both sectors. However, because services compete with the urban population for scarce land, labor productivity will grow slower in services than in manufacturing. Services become more expensive, and the aggregate price level becomes higher with development.

Our model rests on three key observations.

First, most people in the world live in cities, and hence the analysis of price dispersion across cities can potentially inform our analysis of price dispersion across countries. In 2005, around 50 percent of the world population lived in cities.⁴ The overall population density of the earth is rather low, 43 people per square kilometer. However, population is very clustered, so the average person lives in an area with a population density of 5,700 people per square kilometer.⁵ This suggests that the scarcity of land, which has the potential to explain price differences across cities, can also be partly responsible for price differences across countries.

Second, whether a good is tradable or not has implications not only for cross-country trade but also for within-country trade, and hence the location of production within the country. There are many technological differences between tradable sectors such as manufacturing, and non-tradable sectors such as services. The fundamental difference is that non-tradable goods or services have to be produced at the location of their consumption, whereas tradable goods and services can be produced far away from consumers, where land prices may be cheaper. Figure *** shows a map of U.S. counties, displaying the *share* of employment in tradable sectors.⁶ Darker colors represent a higher share in tradable sectors. We see that tradable sectors locate farther away from population centers. This is confirmed in Figure ***, as well, which plots the fraction of employment in traded sectors against the population density of the county. The implication of this observation is that tradable prices are less sensitive to the price of land than non-tradable sectors.

Third, the scarcity of land can be an important constraint to development in at least one sector of the economy: residential housing. Development is usually accompanied by an increase in total factor productivity.⁷ Rich countries can produce more output with the same amount of inputs. This also holds for non-reproducible factors, such as land; after all,

³All data retrieved from <http://bestplaces.net> on January 30, 2008.

⁴World Development Indicators.

⁵We used high-resolution population density data from LandScan to calculate, for the average person on earth, the number of people living within the same 30 arc second by 30 arc second area. (This is basically a population-weighted population density.)

⁶***description

⁷Hall and Jones (1996) find that most of the output per worker differences across countries can be accounted for by differences in TFP.

retail trade has seen one of the highest productivity increase in the past X years in the U.S., where sales per square feet went up from X to Y.⁸ However, land remains a constraining factor for residential housing. Increased productivity in construction can lead to taller and higher-quality buildings, but the “footprint” of a typical household does not decline with development. Davis and Heathcote (2007) estimate that the price of residential land in the U.S. has more than tripled between 1975 and 2005. During this same period, the share of land in the value of a home has increased from 35 percent to 45 percent, suggesting that growth does not bring about a substitution away from residential land.⁹ This implies that land rents increase with development. Together with the previous observation, we have ***

These empirical patterns discipline our modeling choices.

Land rents increase with development because of the scarcity of land available for housing. In the model,

Model results.

The key implication of the model is that relationship between income and the price of non-tradables is stronger among highly urbanized countries than among rural countries. This is because land available for non-tradable production is more scarce in urbanized countries and ***. Empirically, this relationship is strongly borne out in the data. For a cross section of 187 countries in 2005, the ***. This result is robust to alternative definitions of urbanization (population density, the spatial clustering of population).

We emphasize that the above effect is distinct from the direct effect of development on urbanization. Richer countries tend to be more urban; more urban countries, in turn, tend to have higher prices. Even though this effect is also fully consistent with our model, we would like to test the robustness of the *interaction* effect of urbanization and development. We instrument urbanization rates by the geographic features within the country, such as the variation elevation, slope of the terrain (cite Puga), distance to coastline and rivers, and latitude. [***TO BE WRITTEN]

Other, more convoluted testable implications.

Our results do not overturn previous explanations proposed in international macroeconomics or the regional economics literature. Instead, we view them as complementing both types of explanations and potentially bridging the gap between the two. The non-reproducible nature of residential land, together with the optimal location of industries, provides a microfoundation for the *assumption* made by Balassa and Samuelson that non-tradable industries face slower productivity growth.

We then look at how urbanization evolved in the post-world period. Because urbanization was lower, the model predicts the Balassa-Samuelson effect to be weaker. Check in 1960 data. Extrapolating urbanization trends, when did the B-S effect first appear? Our model may provide an explanation for why the Balassa-Samuelson effect is a relatively recent phenomenon.***CHECK. PUT IN NUMBERS. CITE TAYLOR.

⁸Thanks, in a large extent, to the success of Wal Mart.

⁹Also see Davis and Palumbo (2008).

2 A model of housing, development, and relative prices

We introduce our model in three steps. First, we construct the model of housing in a single location and a single production sector (manufacturing). This already enables us to analyze the effect of development on land rents. Then we introduce two locations and analyze the spatial distribution of manufacturing and land rents. Finally, we introduce a nontraded sector in both locations.

In the current version, we are holding several key variables fixed. Development is taken to be exogenous, represented by a Hicks-neutral total factor productivity shifter. Urbanization is also taken as given, modeled as the exogenous spatial allocation of workers across locations. We later wish to analyze the effects of endogenous urbanization.

2.1 A single location

For now we assume a single consumption good, produced using the following technology:

$$x = AF(n, l),$$

where x is output, A is a Hicks-neutral productivity shifter (and the engine of development), $F()$ is a standard CRS production function (quasi-concave and satisfying the Inada conditions), n is the number of workers, and l is the amount of land.

There is another sector, which converts land into residential housing services:

$$h = l_h,$$

where h is service flow from housing and l_h is land devoted to residential investment.

The key assumption regarding the housing sector is that there is no technical progress. This implies that the demand for residential land will increase with development. The scarcity of land will hence constrain the productivity of the other sectors. As we will see, one has to shut down technical progress somewhere so that development leads to changes in relative prices. We believe it is reasonable to assume that the same square footage of housing leads to similar housing services in rich and poor countries, or, at least, that the potential differences are smaller than in any of the productive sectors.

People consume the final good and housing, and their demand is represented by a homothetic, quasi-concave utility function satisfying the Inada conditions:

$$u(x, h).$$

The economy has L amount of land and N workers. Each worker supply a unit of labor inelastically. All land is owned by a lord, who then redistributes the rental income to each worker in the form of lump-sum transfers. Total income of a household is then the sum of wage income and this transfer.

In equilibrium there is full employment of both land and workers. Since workers can only be employed in manufacturing, the only interesting allocation is that of land. Let λ denote the share of land devoted to housing, $\lambda = l_h/L$. Then the supply of housing is λL , and the

supply of the manufacturing good is $AF[N, L(1 - \lambda)]$. Because $F()$ is CRS, we can express the relative supply of manufacturing and housing as

$$\frac{x}{h} = \frac{AF(N/L, 1 - \lambda)}{\lambda}.$$

This is clearly decreasing in λ : the more land is allocated towards housing, the fewer manufacturing goods are produced.

The price of the manufactured good is normalized to 1. The price of housing equals the land rent, which is then determined by the alternative use of land. (The Inada conditions ensure that some manufactured goods will always be produced.) The value marginal product of land in the manufacturing sector is

$$r = AF_L(N, L - L_h) = AF_L(N/L, 1 - \lambda), \quad (1)$$

where the last equality follows from the linear homogenous nature of $F()$. This is clearly increasing in λ : the more land is allocated to housing, the scarcer is land in manufacturing, which raises the marginal product. This equation provides the *marginal rate of transformation* between housing and the manufactured good.

The demand side provides the *marginal rate of substitution* between the two goods. By the homothetic nature of $u()$, the MRS only depends on the relative consumption of x and h ,

$$r = \frac{u_h(x, h)}{u_x(x, h)} = \frac{u_h(x/h, 1)}{u_x(x/h, 1)} = MRS(x/h).$$

This is increasing in x/h . For simplicity, we assume that utility is Cobb-Douglas, so that

$$MRS = \frac{\beta}{1 - \beta} \frac{x}{h},$$

where β is the share of housing in expenditure. This is consistent with the evidence presented in *** that the share of housing expenditure in consumption is constant.

Substituting in the relative supply of x and h ,

$$r = MRS(\lambda) = \frac{\beta}{1 - \beta} \frac{AF(N/L, 1 - \lambda)}{\lambda}. \quad (2)$$

In equilibrium $MRT = MRS$, and equations (1) and (2) can be solved for r and λ .

***figure here

2.1.1 Comparative statics

Equilibrium condition:

$$\lambda F_L(N/L, 1 - \lambda) = \frac{\beta}{1 - \beta} F(N/L, 1 - \lambda),$$

which can be solved for the equilibrium λ . Notice that the equilibrium share of housing does not depend on the level of development, A . This is the result of the Cobb-Douglas utility: since consumers want to spend a constant fraction of their income on housing, they will always devote a constant fraction of land to housing.

Technical progress. We model technical progress by increasing A . If $\sigma = 1$, technical progress does not have an impact on λ . A percentage change in A is directly reflected in a percentage change in r . If $\sigma < 1$, that is, housing and manufacturing are complements in utility, λ increases with development, and the impact of A on rents is more than proportional. ***we may need this for magnification

Population density.

Proposition 1. *Land rents increase with development.*

Proposition 2. *Land rents increase with population density.*

Proposition 3. *Housing increases with population density if and only if land and labor are substitutes in production.*

2.2 Two locations

Now we introduce another location within the country. Both locations have the same endowment of land but they may differ in their population densities, N/L . Without loss of generality, we assume that location 2 has the higher population density, $N_2/L > N_1/L$. We then call this location the “city,” the less populous location the “village.”

The locations interact as follows. Manufacturing goods can be freely transported across locations. People are not allowed to change residence. (We study endogenous urbanization later.) Workers do not commute, so both land and labor are immobile factors.

As before, a lord owns all the land in all locations, and redistributes the rental income for residents in the form of a lump-sum transfer. The budget constraint of the representative household in location i is then

$$x_i + r_i h_i \leq w_i + T.$$

Total spending on manufactured goods (with a price normalized to one) and rent cannot exceed income from wages and the transfer. The balanced-budget condition of the lord is

$$TN = r_1 L + r_2 L.$$

The equilibrium allocation is such that (i) workers and land are fully employed in both locations, (ii) the global market for manufacturing goods clears, (iii) the local markets for housing clear, (iv) firms maximize profits, and (v) households maximize utility given their budget constraint.

In what follows, we focus on equilibria that feature *complete specialization*, that is, manufacturing only active in one location. It is easy to see that this will be the village and no manufacturing takes place within the city. We will later verify the parameter restrictions that require this to be an equilibrium.

In this case, city residents have no wage income, all their income comes from rent transfers. All city land is used for housing, $h_2 = L$. The budget constraint of the city is hence

$$x_2 + r_2 L = \frac{N_2}{N} (r_1 L + r_2 L).$$

Consumption and housing rents have to be covered from transfers that are proportional to total rent in the economy. Letting $\alpha = N_2/N$ denote the fraction of the population that live in the city,

$$r_2 = \frac{\alpha}{1-\alpha}r_1 - \frac{x_2}{L}.$$

We construct the equilibrium in the following steps.

1. Take a land share of housing λ in the village. This pins down rent in the village, r_1 , as the marginal product of land in manufacturing. Clearly, r_1 is increasing in λ .

$$r_1 = AF_L(N_1/L, 1 - \lambda)$$

2. It also pins down the supply of the manufactured good x , because the city has no manufacturing. This supply x is decreasing in λ .

$$\frac{x}{L} = AF(N_1/L, 1 - \lambda)$$

3. Given the rent prevailing in the village and the amount of land allocated to housing, we can derive the villagers' demand for manufactured goods, x_1 . This is increasing in λ , both because of the higher rents, and because of a higher housing consumption.

$$\frac{x_1}{L} = \lambda \frac{1-\beta}{\beta} AF_L(N_1/L, 1 - \lambda)$$

4. The supply of the manufactured good available to city-dwellers, $x_2 = x - x_1$ is hence decreasing in λ .

$$\frac{x_2}{L} = AF(N_1/L, 1 - \lambda) - \lambda \frac{1-\beta}{\beta} AF_L(N_1/L, 1 - \lambda) \quad (3)$$

5. The demand for manufactured goods in the city is pinned down by the city rent, r_2 :

$$\frac{x_2}{L} = \frac{1-\beta}{\beta} r_2 \quad (4)$$

6. There is also the budget constraint of the city,

$$\frac{x_2}{L} = \alpha r_1 - (1 - \alpha)r_2 = \alpha AF_L(N_1/L, 1 - \lambda) - (1 - \alpha)r_2. \quad (5)$$

Equations (3), (4) and (5) can be solved for λ , r_2 and x_2 .

***supply and demand of x_2

2.3 Comparative statics

Technical progress. Note first, that because the equations are linear in (A, x_2, r_2) for any λ , the level of development does not have an impact on the allocation of land. Development raises rents in the two locations as well as the demand and supply for manufactured goods proportionally.

Urbanization. We can also look at the effect of urbanization (a higher fraction of people living in the city) on the allocation of land, production, and prices.

Proposition 4. *Urbanization leads to higher manufacturing output, lower village rents, and higher city rents.*

2.4 Introducing services

We introduce services in the simplest way. The final good x is broken up as a Cobb-Douglas composite of manufacturing and services,

$$x = m^\gamma s^{1-\gamma},$$

so that total utility is

$$u(m, s, h) = m^{\gamma(1-\beta)} m^{(1-\gamma)(1-\beta)} h^\beta.$$

Both manufacturing and services are produced using the same technology

$$\begin{aligned} m &= AF(n_m, l_m), \\ s &= AF(n_s, l_s), \end{aligned}$$

and will hence have the same price in a location in which both are produced. We make this assumption to emphasize that none of our results are driven by technological differences. The only difference between manufacturing and services is that while manufactured goods can be freely traded across locations, services have to be consumed in the location in which they were produced.

The budget constraint of the representative household in location i is

$$m_i + p_i s_i + r_i h_i \leq w_i + T,$$

where p_i is the price of services in location i . Note that this is unity if the manufacturing good is produced in location i (because the two industries have the same cost functions), however, it may be greater than 1 if the manufacturing sector is inactive.

The equilibrium allocation is such that (i) workers and land are fully employed in both locations, (ii) the global market for manufacturing goods clears, (iii) the local markets for services and housing clear, (iv) firms maximize profits, and (v) households maximize utility given their budget constraint.

Again, we focus on complete specialization equilibria. In these equilibria, manufacturing only produces in the village, which pins down services prices in the village as $p_1 = 1$. Service prices in the city are higher, $p_2 > 1$.

In constructing the equilibrium, we first observe that, in the village, both services and manufacturing use the same factor proportions. This is implied by the profit maximization of firms that face common wages, rents, and technologies. If λ share of land is allocated to housing, then the total land used in production is $(1 - \lambda)L$. The ratio of workers to land in both industries is

$$\frac{n_{m1}}{l_{m1}} = \frac{n_{s1}}{l_{s1}} = \frac{N_1}{(1 - \lambda)L}.$$

The marginal product of land only depends on the factor proportions so it continues to be

$$r_1 = AF_L(N_1/L, 1 - \lambda),$$

just as in the case without services. Also, because of the identical technologies, we can simply write the total supply of manufacturing and services in the village as

$$y_m + s_1 = ALF(N_1/L, 1 - \lambda),$$

where y_m is the village supply of manufacturing (which may be different from the village demand).

Cobb-Douglas demand implies constant expenditure shares, so it is easy to determine the demand for services. This also equals the supply of services, because services are nontraded.

$$\frac{s_1}{L} = \frac{(1 - \gamma)(1 - \beta)}{\beta} r_1 \frac{h_1}{L} = \frac{(1 - \gamma)(1 - \beta)}{\beta} \lambda AF_L(N_1/L, 1 - \lambda).$$

The total spending on services and manufacturing is also a constant proportion of housing expenditure,

$$\frac{m_1 + s_1}{L} = \frac{(1 - \beta)}{\beta} r_1 \frac{h_1}{L} = \frac{(1 - \beta)}{\beta} \lambda AF_L(N_1/L, 1 - \lambda).$$

The excess supply (net exports) of manufacturing is supply minus demand, $y_m - m_1$. This is simply

$$\frac{y_m - m_1}{L} = \frac{(y_m + s_1) - (m_1 + s_1)}{L} = AF(N_1/L, 1 - \lambda) - \frac{(1 - \beta)}{\beta} \lambda AF_L(N_1/L, 1 - \lambda), \quad (6)$$

the same as in the previous case, (3).

The budget constraint of the city is

$$m_2 + p_2 s_2 + r_2 h_2 = w_2 N_2 + T N_2.$$

Spending on manufactures, services and housing equal the wage income plus transfers. Because the service technology is CRS, services firms make zero profit:

$$p_2 s_2 = w_2 N_2 + r_2 (L - h_2).$$

The total income of the service sector in the city equals the sum of the total city wagebill and the rents paid on land used by services. The budget constraint can then be rewritten as

$$m_2 + r_2 L = T N_2,$$

which is again the same as before. (Spending on manufactures and the total land rent in the city equals the transfer income.) Writing out transfers,

$$\frac{m_2}{L} = \alpha r_1 - (1 - \alpha) r_2 = \alpha AF_L(N_1/L, 1 - \lambda) - (1 - \alpha) r_2. \quad (7)$$

3 Empirical results

In this section, we present some suggestive evidence supporting the main mechanisms of the model.

3.1 Urbanization and prices

To motivate one of the main propositions of the model, we show that higher urbanization increases the influence of per capita income on the price level and the relative non-tradable/tradable prices in standard cross-country regressions.

Two related, but distinct measures we can use to compare the urbanization level of a country are the proportion of urban population and the population density, both available for a large number of countries from the World Development Indicator database. As our main question is the difference in the 'closeness of residents', both measures are imperfect,¹⁰ but we consider the proportion of urban population a better measure, as it can be expected to capture the clustering of the population better than just the average population density in a country. As the definition of urban population is different in every country, however, the cross-country results should be treated with some caution.

Using the proportion of urban population to divide the sample equally to more rural and more urbanized countries, graph *** shows the cross-country relationships of the (log) price levels and the (log) per capita GDP for both groups, using data from the Penn World Table. The graphs shows, in line with our conjecture, that the relationship between income and the price level is stronger among more urbanized countries. It also shows that the comparison is valid as there are reasonable variation in per capita GDP for both groups, even if there are clearly more high-income countries among the more urbanized ones.

To approach the question more formally, we estimated some regressions in the following form

$$\log P = \alpha_1 + \alpha_2 \log Y + \alpha_3 Z + \alpha_4 (Z - \bar{Z})(\log Y - \overline{\log Y}) + \varepsilon, \quad (8)$$

where P refers to the various price measures, Y is the measure of GDP, and Z -s are the different urbanization measures. Our main interest is the (demeaned) cross-term, as it can be expected to capture how the level of urbanization influences the relationship between the output and the price measures.

Our main price measure is the comparable price level available for over 180 countries from the Penn World Table.¹¹ To check the robustness of the results, we also calculated a non-tradable/tradable relative price index (P_{NT}/P_T) using the basic heading level price data for consumption for the 1996 benchmark ICP countries. This relative price is more in line with the Balassa-Samuelson proposition implying that the main reason of the price level difference is the relatively more expensive non-tradable prices. The measure, however, can

¹⁰Population weighted population-density would be a measure which would express the best the scarcity of land in a country. We are about to develop this measure using the detailed population and land area measures of the LandScan database.

¹¹The sample extends the 115 benchmark countries, that directly participated the 1996 round of the International Comparison Project(ICP), by using somewhat less reliable price data and other observable variables. Our results are robust to restricting the sample to the benchmark countries.

be expected to be less reliable than the aggregate price level as the non-tradable prices, like services, tend to be less comparable internationally than tradable prices,

Table 1: Balassa-Samuelson regressions with urbanization (Robust standard errors are in parentheses)

Explanatory variables	Dependent variables			
	$\log P$		$\log(P_{NT}/P_T)$	
$\log Y$	0.25 (0.05)	0.34 (0.03)	0.28 (0.07)	0.33 (0.05)
urban	0.61 (0.21)		0.12 (0.34)	
urbanX	0.38 (0.11)		0.43 (0.19)	
$\log(\text{density})$		-0.02 (0.02)		-0.02 0.04
densityX		0.05 (0.02)		0.01 0.02
constant	1.13 (0.31)	0.82 (0.26)	-2.96 0.50	-3.13 (0.41)
No. of obs.	186	183	113	113
R^2	0.50	0.46	0.38	0.34

Table 1 implies that the regressions mostly support our conjectures. The first two columns use the (log) price levels as dependent variables, while the last two uses the (log) relative non-tradable/tradable prices. For both dependent variables, we present two regressions with the two urbanization measures. 'Urban' is the proportion of urban population, and 'density' is the population density; and 'urbanX' and 'densityX' are the demeaned output-urbanization cross terms.

In line with well known previous results, higher per capita output increases both the price level and the relative prices across countries, but this relationship is significantly influenced by the level of urbanization, even after controlling for direct price effects of the urbanization. Looking at the first regression, we have found that 1% higher per capita output increases the price level by 0.25% for average level of urbanization (54%), and 1% higher urbanization has a further 0.61% direct effect. The cross effect, however, implies that the effect of output on the prices would be close to 0 under the (extreme) case of 0 urbanization ($0.25 - 0.54 \cdot 0.38 = 0.05$) and would be over 0.4 in case of full urbanization. The effects using the relative prices are similar supporting the conjecture that the price differences are the result of different non-tradable prices. Using the population density as explanatory variable still supports the effect using the price level, though it turns out to be insignificant for our sample in explaining the cross country relative price differences.

3.2 Population

The model conjectures that this effect comes from the facts that, on the one hand, non-tradable products and services need to be produced close to the population and thereby needs to compete for land with residents, and, on the other hand, richer residents drive up the prices of the land. In order to provide some evidence to these mechanisms, we look at US data.