IV. The Costs of Smallness

This chapter is the heart of the report. It takes the various cost data collected, and wherever there are enough of them and they are consistent enough we seek to estimate their relationship with country size. As noted in section II.2, we use a regression approach for the continuous variables and both simple contingency tables and ordered logit models for the categorical ones. Each sub-section discusses the data used and concludes by estimating the cost disadvantages of smallness. For illustrative purposes we present the latter for three exemplar countries located 4th, 18th and 36th in our ranking by size, expressing their percentage disadvantage relative to the median country, ranked 46th. To make the examples concrete, they correspond to the populations of

micro economies	Anguilla	12,13 thousand
very small	Vanuatu	197 thousand
small	Singapore	4,018 thousand
median	Hungary	10,022 thousand

1. Airfreight

<u>Data Sources</u>

The main source for GDP and Population was the World Development Indicators 2002, although other sources were also used to complement this information (Global Business Cost Survey and Asian Development Bank) – see section II.3. Trade data were obtained from the United Nations trade database (Comtrade). For information on air distances, please see the Annex IV.1. Airfreight costs of transporting 100 kilograms of general cargo were obtained from the Survey.

<u>Data concerns</u>

We had doubts concerning 3 survey observations. Marshall Islands had $7.44 typed in for airfreight cost to New York. This was clearly a mistyping and we have multiplied for 100. The value $744 is consistent with costs to London and Tokyo, as well as with those of other countries in the region to New York. Mozambique had typed costs of $1.66, $2.16 and $1.70 to London, Tokyo and New York, respectively. We were unsure if we should multiply this by 100 or 1000. Supporting the first case, the data may refer to 1 Kg of general cargo (or a typing error similar to Marshall Islands). On the other hand, the dot may be marking the thousand units (since in some countries the comma is the decimal indicator and the dot the thousands unit, as the seafreight costs for Mozambique show). If this was the case we should multiply by 1000. There are no data for inbound costs to help us determine which was the mistake. However, comparing the values with the neighbouring countries it seens clear that Mozambique would be a huge outlier if we multiplied by 1000. Thus, the new values are taken as $166, $216 and $170, respectively.

Table 4.1 Airfreight Costs in SADC Countries

	GDP (US$, millions)	Airfreight cost **to** London ($)	Airfreight cost **to** Tokyo ($)	Airfreight cost **to** N.Y. ($)	Airfreight cost **from** London ($)	Airfreight cost **from** Tokyo ($)	Airfreight cost **from** N. Y. ($)	Pop ('000)
Seychelles	613.60	575.00	345.00	465.00	987.00	1601.00	1340.00	81.23
Swaziland	1478.30	225.00	275.00	245.00	308.82	1369.97	300.00	1045.00
Mauritius	4381.40	238.00	330.00	316.00	1990.00	1725.00	1635.00	1186.10
Botswana	5285.20	340.00	504.00	488.00	300.00	550.00	400.00	1602.00
Namibia	3479.20	296.48	320.38	349.07	1550.48	2023.07	1420.34	1757.00
Lesotho	898.95	217.00	257.00	252.00	324.06	1340.56	300.00	2035.00
Zambia	2910.80	398.00	830.00	538.00	540.00	1998.55	1964.00	10089.00
Malawi	1696.60	397.00	470.00	438.00	875.00	2030.00	1340.00	10311.00
Zimbabwe	7392.00	505.00	850.00	530.00	775.00	2000.00	1346.00	12627.00
Mozambique	3753.90	166.00	216.00	170.00	N/A	N/A	N/A	17691.00
Tanzania	9027.50	382.00	762.00	590.00	835.00	1930.00	1375.00	33696.00
South Africa	125890.00	175.00	220.00	235.00	235.29	547.99	295.00	42801.00

Finally, Niue had only two valid observations in the Transportation Costs section in the questionnaire. These are airfreight costs for London ($500.67) and New York ($459), although there is a note stating:

> *Niue is served by a very small aircraft with a maximum passenger carrying capacity of 25 people. This does not allow for carriage of airfreight. Even to the extent that mail delivery is also delayed for weeks at times. It is remote and isolated and is served only through one airline through the Kingdom of Tonga.*

Notwithstanding, we have decided to use these values.

For Vanuatu, the outbound costs were given in foreign currency (Sterling Pounds and Japanese Yen). We used the respective exchange rates to convert them to US dollars.

We also noted that airfreight costs for Antigua & Barbuda, Barbados, Dominica, Grenada and St. Vincent & the Grenadines were potentially inconsistent with the rest of the sample since one of the questions in their questionnaires has been transcribed incorrectly (see Annex IV.1). From inspection it appears that for Grenada and St. Vincent & the Grenadines respondents had answered the (incorrect) question asked rather than the (correct) implicit one, so the Business Cost Survey answers were replaced by those from the following sources:

- Amerijet for Grenada (provided by Mr Junior Mahon)
- Export Development Unit (EDU) for St. Vincent and the Grenadines (provided by Mr Elroy Turner)

In fact, St. Vincent & the Grenadines and Grenada reported the same costs as each other for all destinations.

In the process of obtaining these data, more information was acquired on other Caribbean countries, which were not identical to the survey information. The reason was presumably the discrepancy between different sources, but it serves to highlight, again, the large amount of noise in the survey data. In order to avoid as much as possible the inconsistency of using different sources for outbound and inbound costs (since the new estimates were only available for airfreight outbound costs) we decided not to alter the values from the surveys on the basis of the new data.

The values in table 4.2 were obtained from the Export Development Unit (EDU), and correspond to outbound airfreight costs. In brackets, we have the values from the Global Business Cost Surveys.

Table 4.2 Cost of Airfreight 100 Kilos (US$) - Outbound

Country	**New York**		**London**		**Japan**	
Antigua	187.00	(207.00)	363.00	(317.00)	423.00	(379.00)
Dominica	210.00	(380.00)	364.00	(599.00)	548.00	(921.00)
St. Kitts	228.00	(1654.00)	412.00	(1856.00)	466.00	(2386.00)
St. Lucia	248.00	(-)	644.00	(-)	714.00	(-)
St. Vincent	**239.00**	(754.00)	**351.00**	(1360.00)	**454.00**	(1980.00)
Barbados	169.00	(86.00)	184.00	(90.00)	408.00	(N/A)
Anguilla	255.00	(312.83)	409.00	(422.78)	471.00	(484.79)

Source: Export Development Unit
(Business Cost Survey in brackets)

There is a huge difference between the two sources for St. Kitts & Nevis. However, inbound airfreight costs were the same as those for outbound costs in the Business Cost Survey, which suggests a little consistency. Therefore, considering the data available, sticking with the survey data seemed to be the most coherent approach. In fact, however, in the regression analysis St. Kitts & Nevis had to be dropped once as an outlier.

Table 4.3 Cost of Airfreight 100 Kilos (US$) - Inbound

Country	**New York**	**London**	**Japan**
Antigua	207.00	317.00	379.00
Dominica	380.00	599.00	921.00
St. Kitts	1654.00	1856.00	2386.00
St. Vincent	416.00	719.00	1029.00
Barbados	86.00	90.00	N/A
Anguilla	312.83	422.78	484.79
Grenada	264.00	719.00	966.00

Source: Business Cost Survey

Table 4.4 Missing Observations in Airfreight Costs

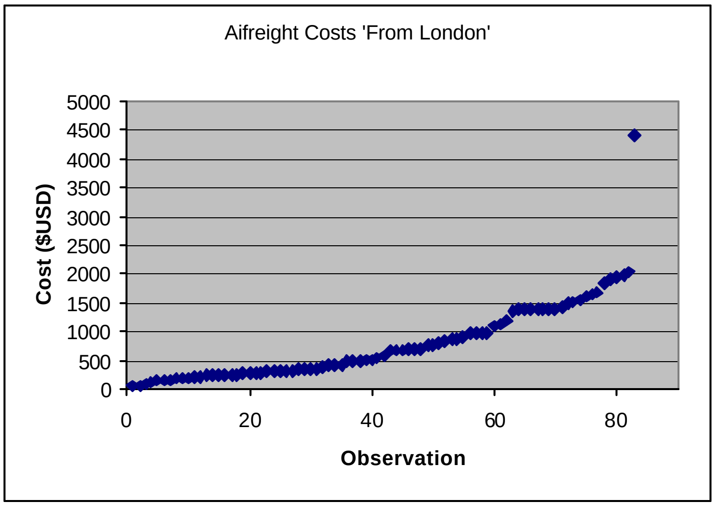
	to London	**from London**	**to Tokyo**	**from Tokyo**	**to NY**	**from NY**
Missing obs.	UK	Cook Is. Guyana Jamaica Kiribati Mozambique Niue Samoa Uganda UK	Barbados Cameroon Japan Niue Uganda	Barbados Cook Is. Guyana Jamaica Japan Kiribati Mozambiq. Niue Uganda	Cameroon Cook Is. Uganda US	Cook Is. Jamaica Kiribati Mozambiq. Niue Samoa Tonga Uganda US Vanuatu

Descriptives

Data plots allowed the detection of possible outliers in the dataset. However, observations were not eliminated solely on the basis of the plots. All outliers were rechecked in the original questionnaire and compared with the other destinations with a view to identifying and correcting obvious mistakes.

The outlier presented in the figure below is Taiwan. We test later if the inclusion of this country affected the results, and the conclusion is "not very much" (partly because the use of logs attenuates its impact: It is 'only' about twice the size of the next largest observation). Greece also presented unexpected values, with the most expensive outbound airfreight and cheapest inbound airfreight being to Tokyo.

Figure 4.1 Airfreight Costs 'From London'



Correlations

The correlations between inbound and outbound costs were analysed as a consistency check. Below is a representative graph - the others look similar.

Figure 4.2 Airfreight Costs: 'From' vs. 'To'

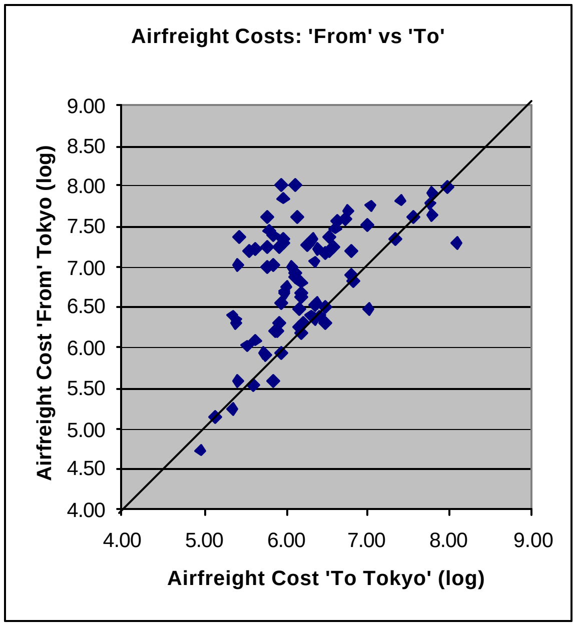


Table 4.5 explores the comparison further. It is clear that most countries have higher inbound than outbound costs.

Table 4.5 Inbound vs. Outbound Airfreight Costs

	London	Tokyo	NY
Correlation coefficient	0.454	0.527	0.675
To > From (%)	28.6	10.9	28.1
To = From (%)	9.5	8.4	8.5
To < From (%)	61.9	80.7	63.4

Analysis of Size

Due to the very high dispersion of Size across our sample and also the view that size effects mainly act proportionally, size variables are logged throughout our analysis. In addition, Hummels (1999)[1] finds a log-linear relationship most useful when modelling the effect of distance on freight costs and we build on that. Thus, for airfreight costs, our initial equation is:

$$\text{Ln(Airfreight Costs)} = \alpha_0 + \alpha_1\text{*Ln(Air Distance)} + \alpha_2\text{*Ln(Size)}$$

The Size measures tried at this stage were GDP (current US$ in millions), Population (in thousands), Total Trade Volume (between the two countries, millions US$) and either Export or Import flows (depending on the direction of the costs, millions US$). As the trade variables performed poorly, and are endogenous anyway, we continued

[1] 'Toward a Geography of Trade Costs ?', University of Chicago (page 10).

the analysis using only GDP and Population as proxies for size, and, as indicated in chapter 2, rapidly converged on population as the correct size indicator.

A second equation was considered to explore any non-linearities in the effects of size. It had the following form,

$$Ln(AirfreightCosts) = \alpha_0 + \alpha_1*Ln(AirDistance) + \alpha_2*Ln(Size) + \alpha_3*[Ln(Size)]^2$$

In some equations it was necessary to delete observations that constituted outliers (i.e., that caused the failure of the test for the normality of residuals). Ecuador, Greece, Nigeria and St. Kitts & Nevis presented such problems for the airfreight regressions 'to Tokyo' and 'to NY'.

The sample of each regression varies since there are different missing values for different airfreight costs.

Results

The results using Population to represent size are reproduced below. The t-statistics are in italics:

Table 4.6 Results for Airfreight Costs

	to London	from London	to Tokyo[1]	from Tokyo[2]	to NY[3]	from NY
constant	**4.876**	**3.945**	**4.302**	**3.702**	**2.349**	**2.284**
	5.103	*3.394*	*4.153*	*4.197*	*2.594*	*1.585*
LnDistance	**0.248**	**0.369**	**0.336**	**0.372**	**0.499**	**0.533**
	3.056	*4.002*	*2.888*	*3.495*	*5.452*	*3.602*
LnPop	**-0.281**	**-0.189**	**-0.326**	**-0.057**	**-0.278**	**-0.211**
	-2.265	*-1.122*	*-2.819*	*-0.421*	*-3.076*	*-1.272*
LnPop2	**0.018**	**0.010**	**0.020**	**0.004**	**0.019**	**0.014**
	2.216	*0.999*	*2.723*	*0.508*	*3.208*	*1.325*
F-test (pop)	2.5648[.083]	0.76963[.467]	4.0068[.022]	0.4687[.791]	5.1788[.008]	0.8858[.416]
R-squared	0.19	0.23	0.14	0.09	0.36	0.17
obs.	91	83	84	80	84	82

[1] Ecuador (+ residual), Greece (+) and Nigeria (+) were excluded due to normality problems.
[2] Malaysia (-), Taiwan (-) and Thailand (-) were excluded due to normality problems. The t-statistics use the White's adjusted standard errors to overcome heteroscedasticity. Furthermore, we used a Wald test (with the White's adjusted standard errors) to test the joint significance of Pop and Pop2.
[3] Ecuador (+), Greece (+), Nigeria (+) and St. Kitts (+) were excluded.

Four simple diagnostic tests were used to evaluate the regressions here and subsequently (see Box IV.1). These are all passed at 5% (i.e., the null hypothesis of a well-specified equation was not rejected) except where noted.

We have also used F-tests to access the joint significance of the size variables. The null hypothesis is that both coefficients (on size and size squared) are equal to zero. The rejection of this hypothesis means that the size variables are jointly significant.

As we can see from the results, Distance is always significant and its coefficients range from 0.25 to 0.53. For Population, the coefficient varies between -0.28 and -0.33 (where significant). The positive (and mostly significant) coefficient on the quadratic form of Population indicates a ‘U-shaped’ relationship between the country size and airfreight costs. In fact, the non-linearity appears to be statistically significant only for the ‘to’ equations, and if we exclude it, population appears to exert no effect on the results at all. In later analysis (chapter V) we accept the latter conclusion of “no effect”, but for now we will continue to permit a quadratic size relationship while we explore the robustness of these equations.

Box IV.1 Selected Regression Specification Tests[2]

Normality test

This is the test proposed by Bera and Jarque (1981) for testing the hypothesis that the disturbances have a normal (Gaussian) distribution, and is valid irrespective of whether the regression equation includes an intercept term or not.

Bera, A.K. and C.M. Jarque (1981), 'An Efficient Large-Sample Test for Normality of Observations and Regression Residuals', Australian National University Working Papers in Econometrics, 40, Canberra.

Ramsey's RESET test for functional form misspecification

Ramsey's (1969) RESET test reported in the diagnostic tests table refers to the simple case where only the square of the fitted values is included in the diagnostic regression of the residuals on the regressors and the squares of the fitted values. The test statistic is equal to the t-ratio of the squared fitted values in this extended regression. If it is significant the functional form of the original equations is rejected, although no indication is given of what is to be preferred.

Heteroscedasticity test

This is a simple test of the hypothesis that the disturbances are homoscedastic, namely that the error variances are equal over the sample. It is calculated from the regression of the squared residuals on squared fitted values and tests whether the squared fitted values in this regression are statistically significant.

Serial correlation

Because we always order observations in terms of increasing size serial correlation of the residuals could be indicative of specification problems. The program computes two statistics

- The Durbin-Watson statistic, and the

[2] These tests are provided in the programme suite Microfit, which we have used in this exercise. For preliminary analysis, they provide sufficient information to assess the performance of our various regressions The descriptions here draw on the Microfit manual.

- Lagrange multiplier (LM) statistic of, for example, Breusch and Godfrey (1981).

Breusch T S and Godfrey L G (1981) ‘ A review of recent work on testing for autocorrelation in dynamic simultaneous models’, in Currie D, Nobay R, and Peel D (ed.s) Macroeconomic Analysis: Essays in Macroeconomics and Econometrics, London, Xcroome Helm, pp. 63-105.

Predictive failure test

This is the second test discussed in Chow (1960), and is applicable even if the number of available observations for the test is less than the number of unknown parameters. It asks whether the additional observations come from the same populations as the sample used in the original regression.

Chow's test of the stability of regression coefficients

This is the first test discussed in Chow (1960), and tests the equality of regression coefficients over two sample periods conditional on the equality of error variances.

Chow, G.C. (1960), 'Test of Equality Between Sets of Coefficients in Two Linear Regressions', Econometrica, 28, pp. 591-605.

<u>Sensitivity Tests</u>

The quadratic form raises the question of whether the relationship is symmetric (i.e., could our negative slope for small countries merely arise because the large country sample required a positive slope?). To test this hypothesis, we created thresholds (for the quadratic form on Size) to analyse if the slopes of the lower or upper segments of the curve were significantly different. The general conclusion was that they were both significant and that Wald tests failed to reject the hypothesis of equal coefficients.
The experimental thresholds were:

Table 4.7 Thresholds for GDP and Population

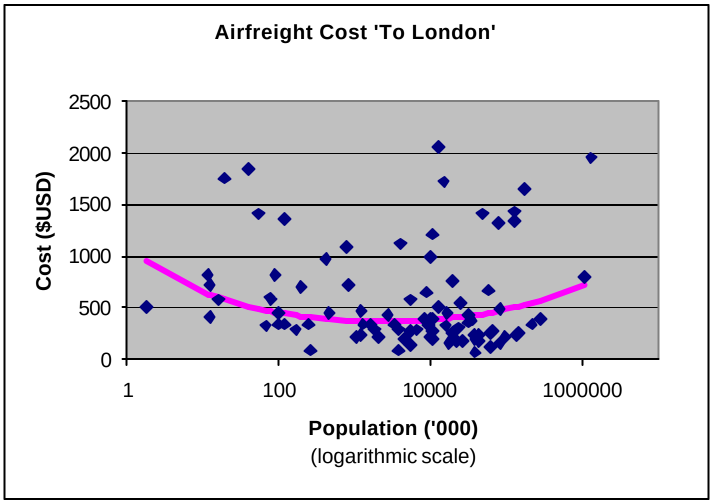
	GDP (millions)			**Population (thousands)**		
Break	5000	15000	30000	1500	3000	6000
Percentile in distribution	38%	49%	51%	33%	36%	45%

We also explored whether China and India could be responsible for the significant curvature. However, they seem to be no different from the rest of the sample.

<u>Cost of Smallness</u>

We now try to summarise the size relationship in a useful way. The turning point of the quadratic equation was calculated and a *ceteris paribus* fitted line estimated. The procedure for the latter was to calculate the fitted cost, evaluated at the sample means of other independent variables, while allowing size to vary according to the country. This gives us an estimate of the cost disadvantage of specific countries.

Figure 4.3 Airfreight Costs 'To London' vs. Population



To produce a set of illustrative results, we have calculated, here and subsequently, the cost disadvantage of three representative countries (which are respectively the 4th, 18th, and 36th - out of 92 - in terms of ascending Population). These countries were chosen in order to give a general view of the cost disadvantage for countries of different size categories (extremely small, very small, and small), relative to the median country. In our sample the median country is Hungary with a population of 10,022 thousand. Thus our cost disadvantage factors essentially compare small economies with an economy of 10 million people with the same GDP pc, distance from major centres, etc. The median is rarely the turning point (minimum for our fitted relationship, so to show its relationship to that point we report the turning point and the country in our sample that is closest to it.

In fact, as Box IV.2 discusses, the resulting estimates offer a somewhat downward view of the cost disadvantages for small countries. But given that the bias leads us to understate cost disadvantages we feel it is appropriate to live with it in this report.

Box IV.2 Biases in estimating the cost disadvantage of size

The cost disadvantage factors discussed throughout the text are calculated from the exponential of the predicted value of the logarithm of costs at the respective population levels. This actually understates the true expected value of the proportionate cost disadvantage for the following reason. The assumption that our equations for ln(costs) have Normally distributed errors entails that the predictions of ln(costs) are Normal and costs, themselves, log-normally distributed. If predicted ln(costs) have distribution $N(m_j, s_j^2)$ for population level j, (where μ_j is the expected value of ln(costs), the true expected value of costs is not exp(μ_j), but exp(μ_j +0.5 s_j). The prediction error from a regression equation grow as one moves further from the mean of the sample data used to fit it. Thus s_j is at a minimum at the mean of the logs of population – about 4.56 million after anti-logging. This corresponds roughly to our small economy, but the exemplar very small economy is 20 times smaller and the micro-economy 335 times smaller! Expressing the disadvantage factors relative to the median – about 2.2 time larger than the mean – means that the expected value of the cost disadvantage of the small economy is slightly overstated, while those of the very small and micro-economies are fairly heavily understated.

Table 4.8 Cost Deviation from the Median Country[3] (%) for Airfreight Costs

Size	Pop ('000)	to London	from London	to Tokyo	from Tokyo	to NY	from NY
Micro	12.13	60.3	*62.1*	85.3	*7.1*	45.2	*37.2*
Very Small	197.00	8.2	*18.9*	15.2	*-0.4*	1.0	*3.2*
Small	4,018.00	-3.1	*1.3*	-2.2	*-1.2*	-4.9	*-3.1*
T point		2,454.20	*12,708.17*	3,463.38	*1,242.65*	1,503.86	*1,873.78*
Closest Country		Jamaica	*Ecuador*	Uruguay	*Gabon*	Botswana	*Namibia*

Here, and throughout, the italicised columns of the cost disadvantages tables are based on statistically insignificant co-efficients. Where the insignificance is crushing and there is no a prior reason to persist with the variable, we eliminate it from our equations. In this case, however, the failure to find a size effect seems surprising a priori, especially given its strength in the 'to' equations, so we present the results at this stage.

Conclusion

We conclude that the size of a country (measured here by its population) can be an important factor explaining its international transport costs. The penalty for smallness is greater for exports than for imports and is of a quite large magnitude. Moreover, although we did not ask about transportation times, small scale almost certainly entails longer delays than normal size, as viable sized transactions are put together. As Hummels shows, time is money when it comes to transporting goods.

Certain aspects of the air transportation costs continue to perplex us. The lack of relationship with size of the inbound costs and the excess of inbound over outbound costs need explanation. The latter may reside in the difficulty of aggregating inbound consignments to an efficient size because they have such different origins and timings. For outbound consignments, by contrast, aggregation is achievable because small economies have low internal communications costs so that the chances of aggregation are better. If this were the case inbound costs could contain an element for returning empty. We do not have the data to test this, and recommend that it be explored in the future.

Annex IV.1

[3] The median country is Hungary, since it ranks as the 46th less populated country. The table reports the population value at the turning point and the country nearest to it in our sample.

I. Air distances

Air distances were calculated using software available at the internet address http://www.wcrl.ars.usda.gov/cec/java/lat-long.htm, which provides great circle distances for most of the cities surveyed for this project. For Belize, Cameroon, Canada, Malawi, Micronesia, South Africa and Vietnam, the distances between the surveyed cities and the three destinations (London, New York and Tokyo) were calculated with the input of coordinates. The coordinates of these cities (that were not shown in the database of the above referred software) were taken from the website http://www.ed-u.com/latitude-and-longitude.htm.

Distances are expressed in Kilometres (km).

II. Transport Costs

Some doubts arise from the questionnaires for the Caribbean.

1. In Jamaica and Trinidad & Tobago analysts used the original survey questionnaires. For the other 7 islands, however, they retyped and reformatted the instrument.

2. The question about airfreight costs for Anguilla and St. Kitts & Nevis is typed '100 kg' (both from the surveyed city and to the surveyed city). For Antigua & Barbuda, Barbados, Dominica, Grenada and St. Vincent & the Grenadines, the questions are typed '*Airfreight cost of transporting **200** kilograms of general cargo from the surveyed city in the surveyed country to* [3 destinations]' and '*Airfreight cost of transporting **100** kilograms of general cargo from the following destinations to the surveyed city in the country*'.

3. Tables:

London	Outward (200 kg)	Inward (100kg)
Antigua & Barbuda	317.47	317.47
Barbados	90	90
Dominica	541	541
Grenada	**1302**	**651**
St. Vincent & Gren.	**1302**	**651**

Tokyo	Outward (200 kg)	Inward (100kg)
Antigua & Barbuda	379.2	379.2
Barbados	NA	NA
Dominica	863	863
Grenada	**1796**	**898**
St. Vincent & Gren.	**1922**	**961**

New York	Outward (200 kg)	Inward (100kg)
Antigua & Barbuda	207.24	207.24
Barbados	86	86
Dominica	322	322
Grenada	**392**	**196**
St. Vincent & Gren.	**696**	**348**

4. The fact that for Grenada and St. Vincent & the Grenadines the outward value is the double the inward, leads us to suspect that the different cargo quantities were taken into consideration (in contrast to the other 3 countries).

IV. The Costs of Smallness
2. Seafreight

Data Sources

The sources for GDP and Population are as stated in the Airfreight section. For sea distances please see the Annex IV.2. Shipping costs of transporting a standard 20ft Full Container Load (FCL) general cargo were obtained from the Survey.

Data concerns

We decided to eliminate a number of European countries (Austria, Czech Republic, France, Germany and Hungary) from the Rotterdam sample since we were not convinced that trade between them and Rotterdam is handled by sea.

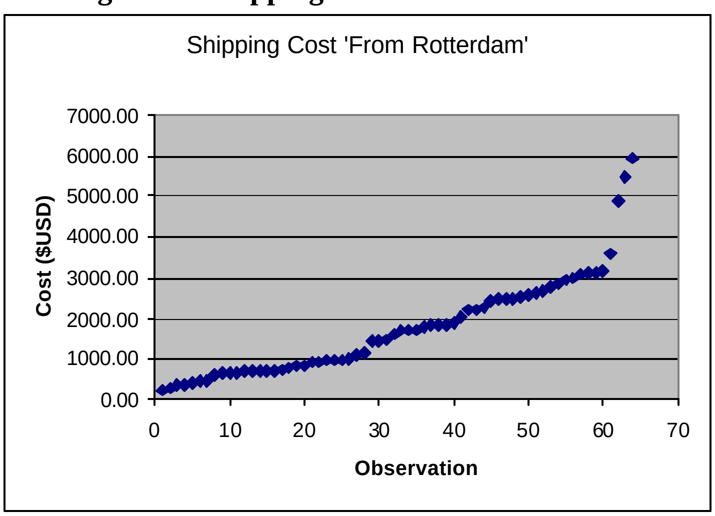
Table 4.9 Missing Observations for Seafreight Costs

	to Rotterdam	from Rotterdam	to Yokohama	from Yokohama	to New York	from New York
Missing obs.	Anguilla Austria Canada Czech Rep. France Germany Hungary Italy Jamaica Marshall Is. Micronesia Nauru Netherlands Niue Palau Poland South Korea Spain	Anguilla Austria Bangladesh Canada Czech Rep. France Germany Hong Kong Hungary India Italy Jamaica Marshall Is. Mexico Micronesia Nauru Netherlands Nigeria	Anguilla Cameroon Canada France Jamaica Japan Nauru Niue Spain Suriname Uganda Venezuela	Anguilla Cameroon Canada France Gabon Japan Mexico Nauru Nigeria Niue Senegal Solomon Is. Spain Tuvalu Uganda	Anguilla Cameroon Canada Cook Is. Fiji Jamaica Kiribati Micronesia Nauru Niue Papua N. G. Samoa Tuvalu Uganda US	Anguilla Canada Cook Is. Fiji Kiribati Mexico Micronesia Nauru Niue Papua N. G. Samoa Suriname Tuvalu Uganda US

	Uganda	Niue				
	Venezuela	Palau				
		Papua N. G.				
		Poland				
		Senegal				
		South Korea				
		Spain				
		Tuvalu				
		Uganda				
		Venezuela				

As previously, we plotted the data on shipping costs. This flagged a number of potential outliers, which we rechecked from the sources and by comparing costs for similar and related countries. Where they were confirmed we decided to treat this problem further ahead. Elsewhere, however, the plots also allowed us to identify and correct a number of typing errors.

Figure 4.4 Shipping Costs 'From Rotterdam'



Two outliers

It was evident that Australia (+) and Finland (-) caused serious normality problems in most of the regressions below. Australia seemed to have extremely high costs in comparison with other countries that have similar distances and are less developed (e.g., Papua New Guinea). Even to Yokohama, Australia is alleged to have the most costly shipping. This just does not seem plausible. On the other hand, Finland has very low shipping costs. 'From [destination]' is exactly the double of 'To [destination]' for all the 3 cases (values 'To' and 'From' from Denmark and Poland are much closer). An illustration of the extent to which they are atypical is given below in table 4.10, which reports studentised residuals from a typical regression.

Table 4.10 Studentised Residuals (u/se)[4]

	to Rotterdam	from Rotterdam	to Yokohama	from Yokohama	to New York	from New York
Australia	2.693	3.210	3.256	3.249	2.901	2.127
Finland	-3.232	-2.156	-3.174	-2.435	-4.087	-2.194
SE of the regression	0.482	0.481	0.524	0.527	0.494	0.608

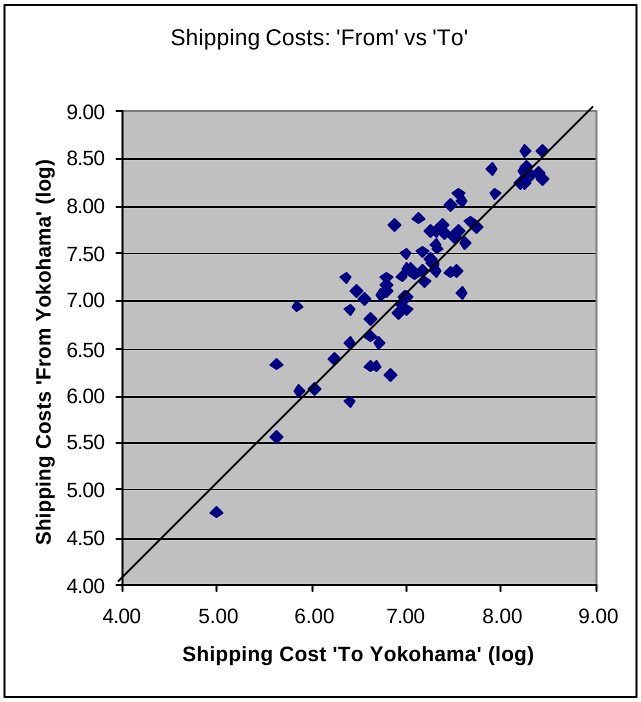
We eliminate these two countries from the shipping freight sample. After this, the only regression with further normality problems was 'From NY'.

Descriptives

Correlations between costs 'from' and 'to' the destination were also calculated. The graphs for Rotterdam and New York are similar to the one depicted below. We checked the cases where there were big differences between 'from' and 'to'.

Figure 4.5 Shipping Costs: 'From' vs. 'To'

[4] These values are calculated based on the regression $Ln(Shipping\ Costs) = \alpha_0 + \alpha_1 * Ln(Sea\ Distance + Land\ Distance) + \alpha_2 * D>500km + \alpha_3 * Ln(Population) + \alpha_4 * [Ln(Population)]^2$



Compared to the airfreight costs, the values for 'from' and 'to' freight rates are much closer. However, countries still typically face higher inbound costs than outbound costs (though to a smaller extent than for airfreight):

Table 4.11 Inbound vs. Outbound Seafreight Costs

	Rotterdam	Yokohama	NY
Correlation coef.	0.881	0.919	0.729
To > From (%)	38.2	18.9	42.7
To = From (%)	7.4	8.1	9.3
To < From (%)	54.4	73.0	48.0

Analysis of Size

Due to missing observations for shipping costs, the sample of each regression varied.

As before, we started by using the following specifications:

$$Ln(Shipping\ Costs) = \alpha_0 + \alpha_1{*}Ln(Sea\ Distance) + \alpha_2{*}Ln(Size) + \alpha_3{*}[Ln(Size)]^2$$

and

$$Ln(Shipping\ Costs) = \alpha_0 + \alpha_1{*}Ln(Sea\ Distance + Land\ Distance) + \alpha_2{*}Ln(Size) + \alpha_3{*}[Ln(Size)]^2$$

Trade variables were not tried on these equations since they are endogenous and performed poorly in the airfreight regressions. For information on distance variables please see Annex IV.2.

Distance

The use of different distance variables did not change the results much, and since container shipping costs include internal transportation, the most coherent procedure is to include land distances (for those countries where the surveyed city is not located on the coast). Thus, land distances in kilometres were converted to nautical miles and added to sea distances. It is clear that this procedure implies that land transportation costs are equal to shipping costs (per mile), which is not true. Therefore, we decided to include a dummy variable to account for this difference. We tried three different specifications:

1 - Use of a dummy for countries with land distance greater than 50 km
2 - Use of a dummy for countries with land distance greater than 500 km
3 - Use of 2 dummies: Dummy for countries with land distance between 50 km and 500 km, and dummy for countries with land distance greater than 500 km

The results indicate that the second specification seems the most relevant. The use of two dummies 'd50-500km' and 'd>500km' only slightly changes the results, with the 'd50-500km' dummy insignificant whilst 'd>500km' is significant in half of the regressions. Since positive land distance is a feature of the sample countries, one would expect to find similar effects from including it for each of the three destination/origin cities. This turned out not to be the case, however, possibly due to the different samples used in each regression. [5]

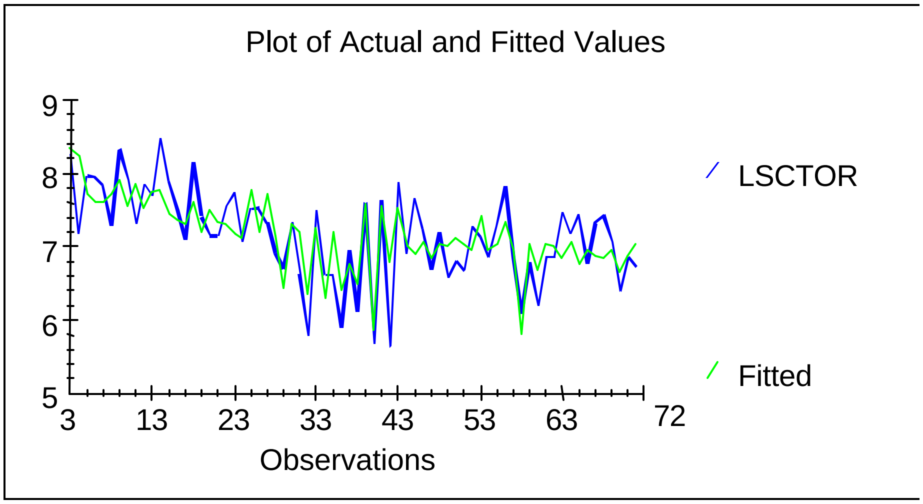
Functional Form

Many of our regressions suggest functional form problems. Several procedures were tried to deal with it:

1 - Estimating a linear rather than log-linear regression, but test results were very poor.
2 - The inclusion of a dummy variable for landlocked countries is not significant in any of the cases.
3 - A specification with 'distance squared' was also considered. Although this procedure seems to solve the functional form problems, the coefficients of population change significantly and the magnitude and sign of some of the distance coefficients were implausible. Moreover, the squares of distance were generally not statistically significant. Thus we do not pursue this route either
4 - Actual and fitted values were plotted to see if there was a systematic feature that could help us to understand what was wrong with the basic specification but we were unable to detect any systematic factors. Below, we have an example of one of these plots:

[5] It may also reflect the logarithmic of the equation, which implies that x km of land distance would be much less significant for Sydney-Rotterdam than, say, Lisbon-Rotterdam.

Figure 4.6 Plot of Actual and Fitted values



5 – A dummy for OECD was included in all regressions, with the justification that OECD countries are significantly different from the rest of the sample in terms of infrastructure and trading institutions (e.g., security), and therefore need to be isolated. The results improved substantially, and most functional form problems disappeared.

Results

Table 4.12 below shows the results:

Table 4.12 Results for Shipping Costs

	to Rotterd.	from Rotterd.	to Yokoh.	from Yok.	to NY[1]	from NY[2]
constant	**6.890**	**7.967**	**4.569**	**3.342**	**7.499**	**8.541**
	8.619	*9.063*	*6.033*	*4.062*	*8.940*	*10.862*
LnDistance	**0.218**	**0.135**	**0.548**	**0.678**	**0.202**	**0.075**
	2.976	*1.731*	*6.312*	*7.372*	*2.290*	*0.889*
LnPop	**-0.290**	**-0.307**	**-0.406**	**-0.316**	**-0.311**	**-0.286**
	-2.765	*-2.422*	*-3.744*	*-2.614*	*-2.576*	*-2.512*
LnPop2	**0.011**	**0.009**	**0.017**	**0.011**	**0.015**	**0.013**
	1.812	*1.1684*	*2.504*	*1.459*	*2.134*	*1.984*
d500	**0.419**	**0.727**	**0.310**	**0.126**	**-0.131**	**0.254**
	2.468	*4.193*	*2.099*	*0.788*	*-0.885*	*1.793*
OECD	**-0.355**	**-0.416**	**-0.291**	**-0.357**	**-0.177**	**-0.279**
	-2.049	*-2.172*	*-2.317*	*-2.683*	*-1.496*	*-1.804*
F-test (pop)	16.601[.000]	29.268[.000]	29.450[.000]	24.973[.000]	5.819[.005]	14.083[.001]
R-squared	0.61	0.70	0.60	0.64	0.29	0.27
obs.	70	62	78	75	75	72

[1] Fails Serial Correlation and Functional Form.

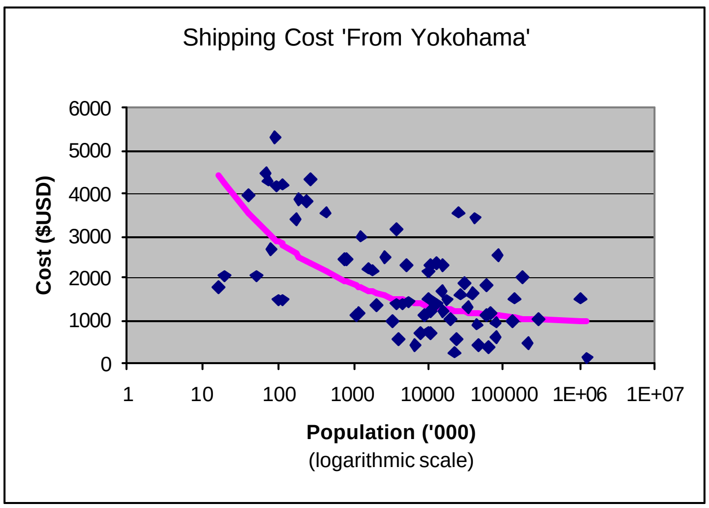
[2] Fails Serial Correlation. Norway (-), Austria (-) and Spain (+) were deleted due to normality problems. The t-statistics use the White's adjusted standard errors to overcome heteroscedasticity. The F-test on population is actually a Wald test (as explained in the airfreight table).

The coefficient on distance varies from 0.08 to 0.55. Population is mostly significant and ranges from -0.29 to -0.41. Again, we used the quadratic form of Population in order to capture a non-linear relationship between Population and shipping costs. The F-tests on Population show us that the inclusion of both POP and POP2 is significant. D500 is significant in the first 3 cases and the last, the positive sign indicating that land transportation is more expensive than by sea.
The negative sign on OECD indicates the lower cost structure that these countries face in comparison with the rest of the sample. As we can see, the regressions for New York have a lower R-squared and look substantially different from those for the other two destinations.

Cost of Smallness

As for airfreight, we set all explanatory variables except size at their mean values (in the estimation sample), and simulate the effects of size. The Population turning point was calculated and a *ceteris paribus* fitted line estimated. This defines the cost disadvantage of the specimen countries.

Figure 4.7 Shipping Costs 'From Yokohama' vs. Population



We can see in table 4.13 that the turning points for seafreight are much higher than those for airfreight, suggesting much higher minimum efficient scale. This, in turn, implies that the size penalty for small states is bigger than in airfreight costs. The graph above reinforces this view.

Table 4.13 Cost Deviation from the Median Country (%) for Seafreight Costs

Size	Pop ('000)	to Rott.	from Rott.	to Yokoh.	From Yokoh.	to NY	From NY
Micro	12.13	195.3	287.4	301.5	251.6	148.3	145.6
Very Small	197.00	67.0	100.1	87.2	85.0	44.4	46.7
Small	4,018.00	9.3	14.6	10.4	11.9	4.5	5.5
T point		530,628.9	*25,534,870.7*	153,457.1	*1,730,007.6*	31,782.4	59,874.1
Closest Country		USA	*N/A*	Brazil	*N/A*	Canada	UK

Conclusion

With the difference that we used dummy variables to correct for differences across regions, the seafreight costs analysis produced much stronger results than the previous one. The penalties are much higher for sea transportation and small countries can have here a big competitive disadvantage when trading with their partners.

Annex IV.2

I. Sea Distances

The main source used to calculate sea distances between ports was software purchased from the internet address http://shipanalysis.com/sud1.html called 'SUDistance v1.3F'. The secondary source used was a National Imagery and Mapping Agency (NIMA) publication called 'Distances between Ports - PUB.151' (2001), downloadable at:
http://pollux.nss.nima.mil/pubs/pubs_j_show_sections.html?dpath=DBP&ptid=5&rid=189.
Nevertheless, many distances where not directly available from these two sources and had to be constructed using one or more of the following procedures:

a) The use of Junction Points along trade routes (e.g. Bishop Rock, Panama Channel, Cape Town, etc.) to calculate sea distances and then sum the lengths.

b) Distance calculation using coordinates where there were no obstacles between the two cities or legs (e.g., The Valley to Bishop Rock).

c) For countries where the surveyed city was not on the coast, we have used major national ports (that may not be the nearest port to the surveyed city). With four exceptions (Colombia, Czech Republic[6], Mexico and Spain) only one port by country was used for calculations. Port statistics were studied to access important ports in Latin America:
http://www.eclac.cl/transporte/perfil/stataec.asp
http://www.eclac.cl/transporte/perfil/stataapa.asp

d) For landlocked countries, we tried to access possible trade corridors in the neighbouring countries. Using various sources like SADC (Southern African Development Community) and online articles, decisions were made after the careful review of relevant documentation.

The main decisions were to use the:
- Beira Port (Mozambique) to serve the cities of Lusaka (Zambia), Harare (Zimbabwe) and Blantyre (Malawi).
- Maputo Port (Mozambique) to serve the cities of Mbabane (Swaziland) and Gaborone (Botswana)
- Durban Port (South Africa) to serve the city of Maseru (Lesotho).

[6] See point d).

- Mombassa Port (Kenya) to serve the city of Kampala (Uganda).
- Trieste Port (Italy) to serve Vienna (Austria), Budapest (Hungary).
- Hamburg Port (Germany) to serve Prague (Czech Republic).

We also calculated land distances from these ports to the surveyed city. For Europe, we used the website http://www.maporama.com/share/, which gives road distances between the main cities in Europe. For the SADC countries, we used free software available form the 'Walvis Bay Corridor Group' (available online at http://www.wbcg.com.na/). Finally, for the Latin American countries, we have used information on road distances provided by travel agents' websites.

Distances are expressed in nautical miles (1 kilometre = 0.5398 nautical mile).

IV. The Costs of Smallness
3. Labour Market
3.1 - Nominal Wages in US dollars

Data Sources

The sources for GDP, GDP PPP (Power Purchasing Parity) and Population were the World Development Indicators 2002, Global Business Cost Survey and Asian Development Bank. Information on wages was taken from the surveys.
With the exception of very few cases for Bank Clerk/Teller and Bank Managers, we had information on wages for all countries.

Data concerns

The survey page regarding the labour market for Namibia had been reformatted. The information provided included minimum and maximum values, as opposed to 'typical' values provided by other countries. We calculated averages. In addition, this survey did not distinguishing between local and foreign banks for Tellers and Managers, and in the absence of any indication about which applied, we excluded this observation.

4.14 Missing Observations for Wages

	Local BC	Foreign BC	Local BM	Foreign BM
Missing obs.	Botswana Kiribati Namibia Niue Papua N. G.	Anguilla Namibia Nauru Tuvalu	Botswana Kiribati Namibia Niue Papua N. G.	Anguilla Namibia Nauru Tuvalu Vanuatu

BC stands for Bank Clerk/Teller and BM for Bank Manager

Descriptives

Plots were used to identify possible outliers and typing errors. The plots of wages against population revealed a great deal of dispersion caused by the inclusion of OECD countries and the three high-income Asian countries (Singapore, Hong-Kong and Taiwan). To maintain a degree of sample homogeneity, the decision was to exclude these economies since they represent completely different realities in terms of labour markets from those we are trying to study[7]. The following graphs illustrate the range, with the later ide ntifying the regions of the sample countries.

Figure 4.8 Kitchen Porter Wage vs. Population

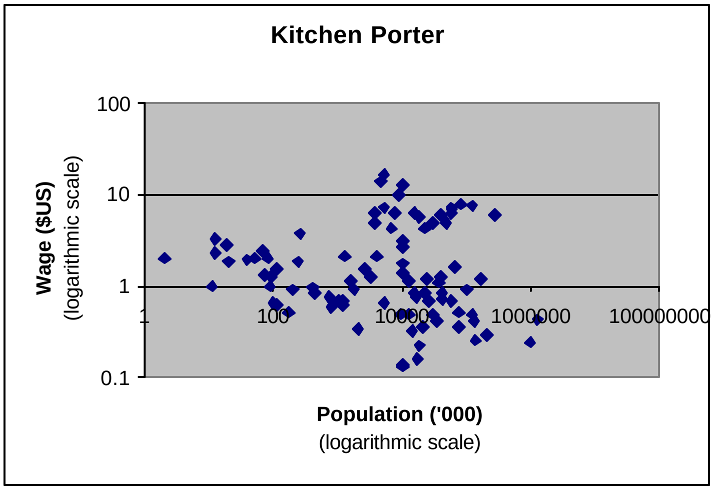
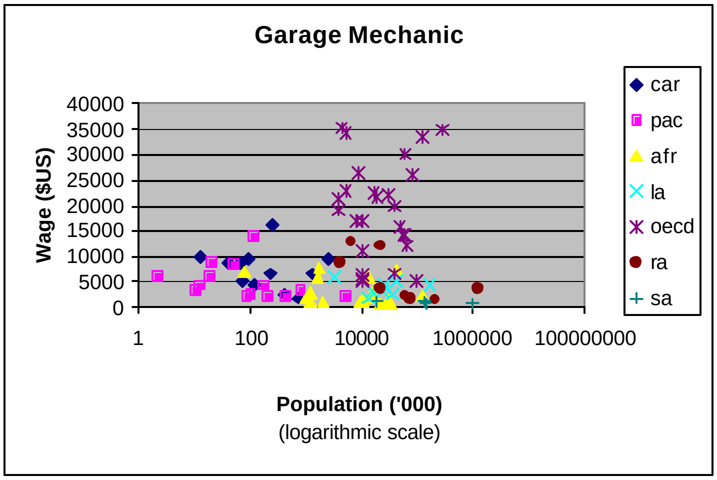


Figure 4.9 Garage Mechanic Wage vs. Population



Analysis of Size

The following specification was used:

[7] For transportation costs one can imagine the same model applying to developed and developing countries, but for labour markets, with all their institutional subtlety, that view is much less persuasive.

$$Ln(Wage) = \alpha_0 + \alpha_1 * Ln(GDP\ per\ capita) + \alpha_2 * Ln(Population)? + \alpha_3 * Benefits$$

One would expect GDP per capita to have a positive sign since it measures the living standards of the country. The effect of Benefits may be positive (in the case that they are complementary to wages, being determined as part of the same real wage process) or, more likely, negative (if they are substitutes).

The Benefits variable used here refers to non-monetary benefits measured as the sum of different types of benefits stated in the Global Business Cost Surveys. Each of 13 classes of benefits was scored 1 for 'present' and 0 for 'absent'. The summary is merely the mean of these scores (i.e., the sum divided by 13). In fact, however, this variable performed poorly (perhaps because the different effects cancel each other), so it was excluded from our reported results.

In line with the experiments on freight charges, we also used a second specification:

$$Ln(wage) = \alpha_0 + \alpha_1 * Ln(GDP\ per\ capita) + \alpha_2 * Ln(Population)? + \alpha_3 * [Ln(Population)]^2$$

However, none of the squared terms on population proved to be significant, so we do not pursue them here.

Finally, we tested an equation that allows for insularity effects. The fact that some countries in our sample are islands may affect their wages, in the sense that insularity is an obstacle to labour mobility. To account for this issue we experimented with the following dummies (see chapter II.3):

D50 - Island with less than 10 million inhabitants and without any considerable mass of land within 50 km. The exception is New Zealand due to its degree of development. Papua New Guinea was not included because it shares the island with a territory of Indonesia (Irian Jaya, or *West Papua*). Furthermore, Trinidad and Tobago was included in this category because the distance between Port of Spain and Cumana (the most developed port in the Northeast of Venezuela) is greater than 50 km.
D500 and D1500, which have the same definitions but extend the threshold distance, to 500 km and 1500 km respectively.

Table 4.15 Measures of Insularity

	Caribbean	**Pacific**	**Africa**
D50	Anguilla Antigua & Barbuda Barbados Dominica Grenada Jamaica St. Kitts & Nevis St. Vincent & Gren. Trinidad & Tobago	Cook Islands Fiji Marshall Islands Kiribati Micronesia FS Nauru Niue Palau Samoa Solomon Tonga Tuvalu Vanuatu	Mauritius Seychelles

D500 (change to D50)	-Grenada -St.Vincent & Gren. -Trinidad & Tobago		
D1500 (change to D500)	-Rest of Caribbean Islands	-Palau	-Mauritius

<u>Results</u>

Table 4.16 presents the main results:

4.16 Results for Wages

	CW	CO[1]	KP	BCL	BCF	GM	PC[2]	QT	BML[3]	BMF[4]	GRN[5]
constant	**0.290**	**0.024**	**0.151**	**8.509**	**8.555**	**8.304**	**7.957**	**8.396**	**9.906**	**9.997**	**8.545**
	1.332	*0.146*	*0.949*	*36.58*	*35.402*	*40.402*	*37.012*	*47.279*	*32.98*	*30.432*	*46.232*
LnGDPc	**0.525**	**0.489**	**0.452**	**0.257**	**0.229**	**0.461**	**0.543**	**0.532**	**0.298**	**0.291**	**0.515**
	7.559	*6.934*	*8.907*	*3.651*	*3.167*	*7.039*	*7.930*	*9.403*	*3.259*	*2.994*	*8.715*
LnPop	**-0.075**	**-0.054**	**-0.080**	**-0.050**	**-0.040**	**-0.060**	**-0.012**	**-0.047**	**-0.064**	**-0.046**	**-0.071**
	-3.064	*-3.057*	*-4.350*	*-1.937*	*-1.491*	*-2.579*	*-0.477*	*-2.337*	*-2.077*	*-1.281*	*-3.437*
Obs	63	63	63	58	59	63	63	63	58	58	61
R-squared	0.67	0.67	0.76	0.39	0.29	0.63	0.60	0.73	0.35	0.26	0.73

1 t-statistics use the White's adjusted standard errors to overcome heteroscedasticity
2 Fails Serial Correlation
3 t-statistics use the White's adjusted standard errors to overcome heteroscedasticity
4 Fails Serial Correlation
5 Sri Lanka (-) and Zimbabwe (+) were eliminated due to normality problems

<u>Definitions</u>:
CW - Construction Worker
CO - Checkout Operator
KP - Kitchen Porter
BCL - Bank Clerk/Teller in Local Bank
BCF - Bank Clerk/Teller in Foreign Bank
GM - Garage Mechanic
PC - Payroll Clerk
QT - Qualified Teacher
BML - Bank Manager in Local Bank
BMF - Bank Manager in Foreign Bank
GRN - General Registered Nurse

One can see that population is significant in most of the regressions, with elasticities ranging from -0.05 to -0.08 (where significant). GDP per capita is strongly significant for all regressions and varies between 0.45 and 0.54 with the exception of bank related jobs (i.e., Bank Clerk and Bank Manager). In fact, these four cases are obviously different from the others: the R-squares are lower, and so are the coefficients of GDP per capita and Population.

4.17 Results for Wages with D50

	cw	co[1]	kp	bcl	bcf	gm	pc	qt	bml[2]	bmf	grn[3]

constant	**-0.394**	**-0.500**	**-0.014**	**8.431**	**8.309**	**7.855**	**7.470**	**8.249**	**10.137**	**10.069**	**8.185**
	-1.103	*-1.702*	*-0.050*	*21.28*	*20.78*	*22.784*	*20.749*	*27.186*	*20.519*	*18.786*	*26.211*
LnGDPc	**0.506**	**0.475**	**0.448**	**0.255**	**0.222**	**0.448**	**0.529**	**0.528**	**0.306**	**0.294**	**0.506**
	7.509	*6.997*	*8.724*	*3.537*	*3.033*	*6.891*	*7.793*	*9.216*	*3.270*	*2.961*	*8.580*
LnPop	**-0.012**	**-0.006**	**-0.063**	**-0.043**	**-0.017**	**-0.018**	**0.033**	**-0.033**	**-0.085**	**-0.053**	**-0.038**
	-0.336	*-0.212*	*-2.315*	*-1.094*	*-0.439*	*-0.529*	*0.929*	*-1.095*	*-1.805*	*-0.996*	*-1.237*
D50	**0.532**	**0.407**	**0.128**	**0.060**	**0.189**	**0.349**	**0.379**	**0.114**	**-0.177**	**-0.057**	**0.280**
	2.370	*2.236*	*0.750*	*0.246*	*0.773*	*1.611*	*1.672*	*0.598*	*-0.558*	*-0.172*	*1.425*
Obs	63	63	63	58	59	63	63	63	58	58	61
R-squared	0.70	0.69	0.76	0.39	0.30	0.64	0.62	0.73	0.36	0.26	0.74

[1] The t-statistics use the White's adjusted standard errors to overcome het eroscedasticity
[2] The t-statistics use the White's adjusted standard errors to overcome heteroscedasticity
[3] Sri Lanka (+) and Zimbabwe (+) were eliminated due to normality problems

Adding insularity, D50 is significant only for the two first cases and marginally for Payroll Clerks. However, its inclusion renders population insignificant (except for Kitchen Porter, while BML becomes significant). Altering the distance threshold (by replacing D50 with D500 or D1500), allows the Population coefficient to start converging to their previous values and the insularity measure becomes insignificant for all cases.

The insularity effect seems mainly to capture the atypically high levels of wages in the Caribbean, and this is why it tends to undermine the size effects. If separate dummies are included for the Caribbean and the Pacific, they typically show a positive effects with the former larger and more significant than the latter, while the population effect is poorly determined. We choose not to use regional dummies in the equations for this chapter (in order to allow better identification of size effects) and we prefer size over insularity in explaining wages (since the Caribbean and Pacific effects have opposite signs). However, the ‘excessive’ wages of the Caribbean region are worthy of future exploration, and if they can be satisfactorily explained one would have to consider the possibility that our size effects are over-stated.

Sensitivity tests

Differences in nominal dollar wages do not merely reflect living standards but also differences in the cost of living in small countries. If this were the whole explanation, we would be in danger of conceptual double counting in the costs of smallness - once via extra cost of the sort explored elsewhere in this section, and once via their effects on nominal incomes. To allow for the cost of living we add the logarithm of the PPP factor to the estimation equation. [8] As expected it shows a negative effect on nominal wages, suggesting that where prices are higher (the PPP factor is lower), wages are higher. The effect is not particularly well defined, however. The population effect in these equations is typically weaker and less significant than before, but it is still always negative. The problem in these equations is that the PPP factor and size are strongly collinear, as table 4.19 below shows. (We report one regression of PPP on population for each wage type to allow for the slightly different samples). Larger size consistently predicts lower local prices relative to international levels.

Table 4.18 Regressions with GDPpc(PPP)/GDPpc (named PPP)

[8] The PPP factor is measured as GDP (PPP, $)/GDP(current, $). Lower values imply higher current prices relative to the global PPP values used.

	cw	co	kp	bcl	bcf	gm	pc[1]	qt	bml[3]	bmf[4]	grn
constant	**0.4**	**0.09**	**0.249**	**8.625**	**8.727**	**8.35**	**8.132**	**8.412**	**9.925**	**10.102**	**8.593**
	1.805	*0.45*	*1.559*	*36.78*	*35.194*	*39.077*	*39.209*	*45.354*	*31.662*	*29.289*	*44.81*
lgdp_c	**0.445**	**0.441**	**0.381**	**0.172**	**0.126**	**0.427**	**0.415**	**0.52**	**0.284**	**0.225**	**0.48**
	5.552	*6.056*	*6.597*	*2.108*	*1.481*	*5.524*	*5.53*	*7.743*	*2.604*	*1.918*	*6.915*
lpop	**-0.042**	**-0.034**	**-0.048**	**-0.016**	**-0.004**	**-0.046**	**0.041**	**-0.042**	**-0.059**	**-0.023**	**-0.057**
	-1.411	*-1.261*	*-2.258*	*-0.533*	*-0.142*	*-1.593*	*1.484*	*-1.676*	*-1.44*	*-0.522*	*-2.198*
l(PPP)	**-0.401**	**-0.245**	**-0.358**	**-0.415**	**-0.485**	**-0.167**	**-0.643**	**-0.058**	**-0.069**	**-0.316**	**-0.179**
	-1.873	*-1.26*	*-2.32*	*-1.951*	*-2.126*	*-0.811*	*-3.205*	*-0.326*	*-0.242*	*-1.002*	*-0.949*
Obs	63	63	63	58	59	63	63	63	58	58	61
R-squared	0.69	0.68	0.78	0.43	0.34	0.63	0.66	0.73	0.35	0.28	0.74

[1] Fails Serial Correlation
[3] Fails Heteroscedasticity
[4] Fails Serial Correlation

Table 4.19 Regressions of Ln(PPP)=a+b*Ln(Pop)

	cw	co	kp	bcl	bcf	gm	pc	qt	bml	bmf	grn
constant	**-0.095**	**-0.095**	**-0.095**	**-0.135**	**-0.046**	**-0.095**	**-0.095**	**-0.095**	**-0.135**	**-0.071**	**-0.089**
	-0.782	*-0.782*	*-0.782*	*-1.007*	*-0.341*	*-0.782*	*-0.782*	*-0.782*	*-1.007*	*-0.517*	*-0.734*
lpop	**0.118**	**0.118**	**0.118**	**0.122**	**0.113**	**0.118**	**0.118**	**0.118**	**0.122**	**0.115**	**0.116**
	8.152	*8.152*	*8.152*	7.712	*7.102*	*8.152*	*8.152*	*8.152*	7.712	7.205	*7.981*
Obs	63	63	63	58	59	63	63	63	58	58	61
R-squared	0.52	0.52	0.52	0.52	0.47	0.52	0.52	0.52	0.52	0.48	0.52

The collinearity leaves a dilemma as to how to conceptualise the costs of smallness, but practically it makes little difference. Table 4.20 substitutes the estimated PPP/size relationship from table 4.19 into wage equations in table 4.18 and calculates a combined direct and indirect size effect. The resulting coefficients are similar to those presented in table 4.16 above as our main equations, but in every case absolutely larger. Thus our main equations offer a conservative view of the wage gradient with size.

Table 4.20 Population Effects on Nominal Wages

	cw	co	kp	bcl	bcf	gm	pc	qt	bml	bmf	grn
Direct	-0.075	-0.054	-0.080	-0.050	-0.040	-0.060	-0.012	-0.047	-0.064	-0.046	-0.071
Indirect	-0.089	-0.063	-0.090	-0.066	-0.059	-0.066	-0.035	-0.049	-0.067	-0.059	-0.078

Cost of Smallness

Since we have rejected the inclusion of the squared term for population, we do not have turning points, but there are still cost disadvantages to smallness. The figures below are derived from table 4.16.

Table 4.21 Cost Deviation from the Median Country (%) for Wages

Size	Pop ('000)	CW	C0	KP	BCL	BCF	GM	PC	QT	BML	BMF	GRN
Micro	12.13	65.5	43.7	71.1	39.9	*30.8*	49.6	*8.4*	37.1	53.7	*36.2*	61.1
Very Small	197.00	34.3	23.6	36.9	21.7	*17.0*	26.6	*4.8*	20.3	28.6	*19.8*	32.2
Small	4,018.00	7.1	5.1	7.6	4.7	*3.7*	5.6	*1.1*	4.4	6.0	*4.3*	6.7

Conclusion

We conclude that, among developing countries, i.e. excluding OECD, Hong Kong, Taiwan and Singapore (the high-income economies), there is a significant population effect on nominal dollar wages. In the sense that higher nominal wages in the small states adversely influence external competitiveness, business in these states pays a penalty for smallness.

IV. The Costs of Smallness
3. Labour Market
3.2 - Availability of Workers

Data Sources

Data on GDP and Population were obtained from the sources mentioned before. Information on the availability of workers was taken from the Survey.

Descriptives

We have information regarding availability of unskilled, semi-skilled and skilled workers. This information was taken from three 'closed' questions, one for each skill category, in which only one answer should be chosen. The question was:
"*Are unskilled/semi-skilled/skilled workers domestically available or they have to be hired from abroad?*"
The options presented were:

1- Domestically available and there is no need to import them from abroad.
2- There are enough workers domestically available to satisfy most of the demand, however occasionally workers need to be imported from abroad.
3- Workers are available domestically, but many need to be imported from abroad to satisfy demand
4- Some workers are domestically available, but most need to be imported from abroad
5- Few workers are available domestically. The vast majority of workers are imported from abroad

Two countries recorded more than one entry. We checked the remarks in the surveys and decisions were made case-by-case. Where no conclusion was possible we opted for the worse scenario. These were for Botswana for all three categories, and Namibia for semi-skilled labour.

Table 4.22 Data Decisions for Availability of Workers

Country	Question	Answers	Remarks	Decision
Botswana	Unskilled	(1) and (2)	-	(2)
	Semi-skilled	(2) and (3)	-	(3)

	Skilled	(3) and (4)	There is a general shortage of skilled/professional workers in Botswana. (...) more than two thirds of the positions are held by non-citizens.	(4)
Namibia	Semi-skilled	(2) and (3)	-	(3)

Below we present data cross-classified by labour market conditions and size as measured alternatively by population and GDP.

The categories of availability for labour (1-5) correspond to the answers presented above. For the independent data, the standard classes defined in table 2.2 are applied.

The following tables give the cross-tabulations of the need to import workers from abroad and size. We consider both Population and aggregate GDP as indicators of size, but since we are making most of our arguments with Population, we focus mainly on that[9].

Unskilled workers

Table 4.23 Cross Tabulations for Unskilled Workers (Population)

Availability of Workers	Population					
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	Total
1	13	8	9	22	16	68
2	4	2	2	2	2	12
3	2	-	1	5	-	8
4	-	-	2	1	-	3
5	1	-	-	-	-	1
Total	20	10	14	30	18	92

Table 4.24 Cross Tabulations for Unskilled Workers (GDP)

Availability of Workers	GDP					
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	Total
1	13	7	13	15	20	68
2	2	2	2	-	6	12
3	-	1	1	1	5	8
4	-	-	-	2	1	3
5	1	-	-	-	-	1
Total	16	10	16	18	32	92

For unskilled workers it is difficult to detect any population effect. For 90% of countries with population greater than 50 million workers are totally available (domestically), while for the other population categories this percentage is lower (especially categories 1 and 3). Generally, there are no severe shortages of unskilled workers in our sample of countries. Only 12 countries admit regular labour import,

[9] We will briefly explore a more sophisticated model for these categorical data – the ordered logit - but for the present explanatory exercise on relatively few relatively noisy data we believe that simple contingency tables capture the essential story.

and half of them are in countries with population from 10 to 50 million (only four report serious problems).

To conduct any formal tests of statistical significance on these numbers, the classifications need to be collapsed into a much smaller dimensionality. Thus we group the information into 2-by-5 tables and examine Pearson's Chi-square statistic for the null hypothesis of no association. Since this is experimental work we aggregate the values in two different ways. First, we group categories 2 to 5 and test against category 1 alone: this amounts to testing complete availability against the need for any import at all. Second we test categories 1 and 2 against 3-5 which draws the divide at regular importing of workers v non, or only occasional importing. Even with these aggregations the conditions for the Chi-squared test are only barely met, so a degree of caution is called for in interpreting the results.

Table 4.25 Chi-Square Statistics for Unskilled Workers

Type of Grouping	Pearson's Chi Square		
	Value	df	Asymp. Sig. (2 sided)
Pop (1, 2-5)	3.788	4	0.435
Pop (1-2, 3-5)	6.415	4	0.170
GDP (1, 2-5)	3.963	4	0.411
GDP (1-2, 3-5)	2.511	4	0.643

These test statistics easily fail to reject the null hypothesis that there is no significant size (Population or GDP) effect on the availability of unskilled workers.

The ordered logit equations, which determine the effect of population and GDP pc on the chances of a country falling into one of the above categories, confirm that, although smaller economies appear to more likely to suffer labour shortages, there is no significant relationship (at 5%) between size and the availability of unskilled labour. (The co-efficient was negative – smaller countries face higher probabilities of having to import labour – but significant only at 10%).

Semi-skilled workers

Table 4.26 Cross Tabulations for Semi-skilled Workers (Population)

Availability of Workers	Population					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
1	7	5	7	23	17	59
2	10	2	6	7	1	26
3	2	3	1	-	-	6
4	-	-	-	-	-	0
5	1	-	-	-	-	1
Total	20	10	14	30	18	92

Table 4.27 Cross Tabulations for Semi-skilled Workers (Population)

Availability of Workers	GDP					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	

1	**6**	**4**	**11**	**17**	**21**	**59**
2	**7**	**6**	**2**	**1**	**10**	**26**
3	**2**	**-**	**3**	**-**	**1**	**6**
4	**-**	**-**	**-**	**-**	**-**	**0**
5	**1**	**-**	**-**	**-**	**-**	**1**
Total	**16**	**10**	**16**	**18**	**32**	**92**

Only 7 countries report a regular need to import semi-skilled workers from abroad. However, if we look at the percentages of countries that have sufficient labour supply (with no need to import even occasionally) there exists a hint of a population effect, with the percentage falling in smaller states. Only 35% of the countries with lower population completely satisfy their own demand for labour, although, cumulating the first two categories, 85% still need to import only occasionally or never.

We used the same procedure as before to test formally the significance of the hypothesis of no association, and here there is a strong rejection of the null, with small countries needing to import semi-skilled labour more frequently than larger ones. This is true for both aggregations of the data and both concepts of size.

Table 4.28 Chi-Square Statistics for Semi-skilled Workers

Type of Grouping	**Pearson's Chi Square**		
	Value	**df**	**Asymp. Sig. (2 sided)**
Pop (1, 2-5)	**18.701**	**4**	**0.001**
Pop (1-2, 3-5)	**12.644**	**4**	**0.013**
GDP (1, 2-5)	**14.834**	**4**	**0.005**
GDP (1-2, 3-5)	**8.871**	**4**	**0.064**

The ordered logit equation, which determines the effect of population and GDP pc on the chances of a country falling into one of the above categories, confirms that smaller economies are significantly more likely to suffer shortages of semi-skilled labour. The co-efficient on population is strongly significantly negative. To give an indication of the role of population, table 4.29 below gives the population levels for which a country with GDP pc of $10,000 is predicted to lie in each labour availability category. It must be remembered that these are ceteris paribus predictions, holding income constant and ignoring a host of specific factors that would affect any real country. Nevertheless, the table shows that as population falls we should expect countries to report more difficulties over labour availability.

Table 4.29 Ordered Logit for Semi-skilled Workers

Availability of Workers	**Population ('000) [if GDP pc = $10,000]**
1	**Above 1,893**
2	**From 4 to 1,893**
3	**Below 4**
4	

5	

Skilled workers

Table 4.30 Cross Tabulations for Skilled Workers (Population)

Availability of Workers	Population					
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	Total
1	3	-	4	12	13	32
2	9	4	6	13	5	37
3	5	5	4	5	-	19
4	3	1	-	-	-	4
5	-	-	-	-	-	0
Total	20	10	14	30	18	92

Table 4.31 Cross Tabulations for Skilled Workers (GDP)

Availability of Workers	GDP					
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	Total
1	2	1	2	9	18	32
2	5	5	8	8	11	37
3	6	4	5	1	3	19
4	3	-	1	-	-	4
5	-	-	-	-	-	0
Total	16	10	16	18	32	92

These tables are the most striking, although we would look a bit sceptically upon the second population category. Nevertheless, only 15% of the small countries completely satisfy their own labour demand, while 40% of them have a clear need to import labour. If we compare the small states with the third population category and larger, there is a clear difference. Small countries seem to lack enough skilled workers, even though they tend to pay higher salaries.

We complement the tables with the same tests as before.

Table 4.32 Chi-Square Statistics for Skilled Workers

Type of Grouping	Pearson's Chi Square		
	Value	df	Asymp. Sig. (2 sided)
Pop (1, 2-5)	20.505	4	0.000
Pop (1-2, 3-5)	16.140	4	0.003
GDP (1, 2-5)	18.050	4	0.001
GDP (1-2, 3-5)	18.663	4	0.001

The Person Chi-squares are higher for the skilled category than for other labour and strongly reinforce the conclusion above. Small countries clearly have to import skilled labour. This may lie behind the higher wages discovered earlier, but it compounds the disadvantage implied by those wages. Imported labour is much less flexible for business and to the extent that there are external benefits from having skilled citizens (for example in citizenship, the effectiveness of education), small countries miss out.

Table 4.33 Ordered Logit For Skilled Workers

Availability of Workers	Population ('000) [if GDP pc = $10,000]
1	Above 15,803
2	From 55 to 15,803
3	Below 55
4	
5	

Table 4.33 reports the ordered logit exercise for skilled workers. Compared with table 4.29 these results indicate that small countries face an even more serious shortage of skilled than semi-skilled workers.

Conclusion

The data appear to indicate material shortages of skilled workers in small countries (and in some cases also semi-skilled). This may be an important factor behind the higher wages for skilled jobs, although in a perfectly functioning market the wage would rise sufficiently to damp down the excess demand. All told, both wages and labour shortages seem likely to hinder international activities in very small countries. It is difficult to assess its practical importance, but it is worth noting that the survey questions implicitly enquire whether skills are sufficiently available for the current set of activities. These, in turn, are determined by relative endowments via the traditional trade theory effects on the composition of output. Thus labour shortages will have a greater impact on skill-intensive activities than a simple reading of these results would suggest: they will both reduce the extent of such activities and make any remaining activity vulnerable for labour supply.

IV. The Costs of Smallness
3. Labour Market
3.3 – Manufacturing Costs

Data Sources

Data on GDP and Population were obtained from the same sources as above. Information on manufacturing costs was taken from the Survey (although we only had 63 valid observations in all).

Data concerns

One might hope that a question on manufacturing labour costs would pull together the previous results on wages and labour shortages, as well as allow for productivity differences. In fact, however, we are very sceptical of the ability of this question to do so. First, we only have 11 valid observations among the 20 least populated countries, and second, the question 'Manufacturing labour cost per hour for the whole country' is very broad, to say the least.

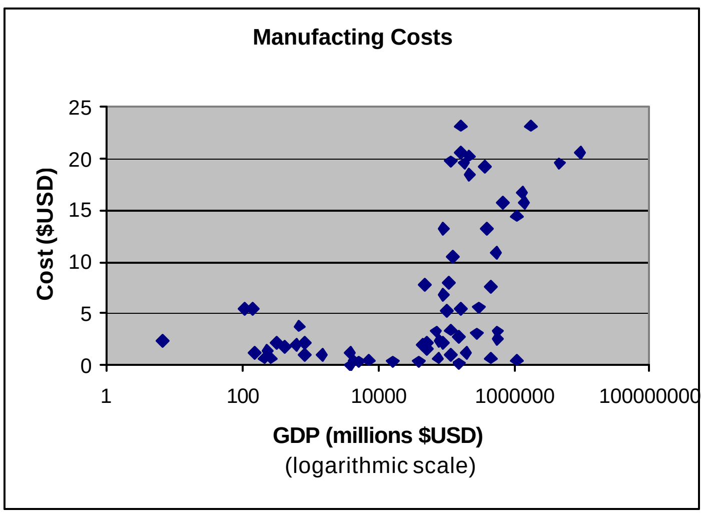
Table 4.34 Missing Observations for Manufacturing Costs

	Manufacturing Costs		
Missing Obs.	Bangladesh	Kenya	St. Vincent & Gr.
	Barbados	Kiribati	Swaziland
	Cameroon	Lesotho	Tanzania
	Cook Islands	Malawi	Trinidad & Tob.
	Cote d'Ivoire	Marshall Is.	Tuvalu
	Dominica	Micronesia	Uganda
	Ecuador	Namibia	Uruguay
	Gabon	Nauru	Vietnam
	Guyana	Pakistan	Zambia
	Jamaica	Senegal	

Descriptives

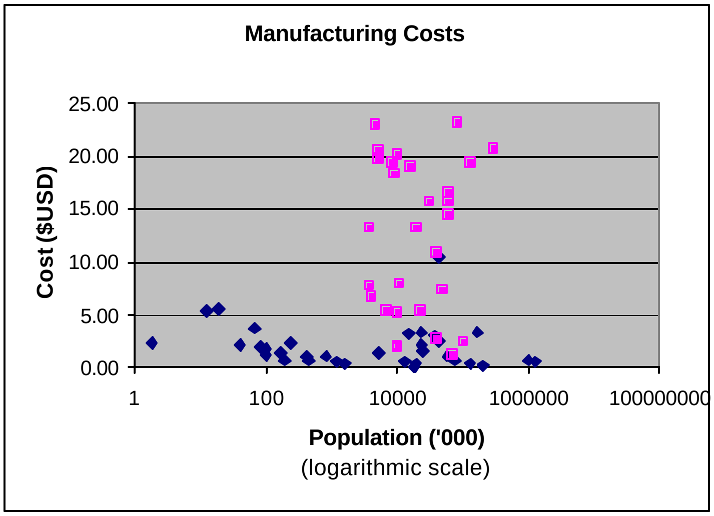
Below we plot manufacturing costs against GDP. We can see that there is a clear break in the data, largely caused by the inclusion of OECD and Asian high-income countries.

Figure 4.11 Manufacturing Costs vs. GDP



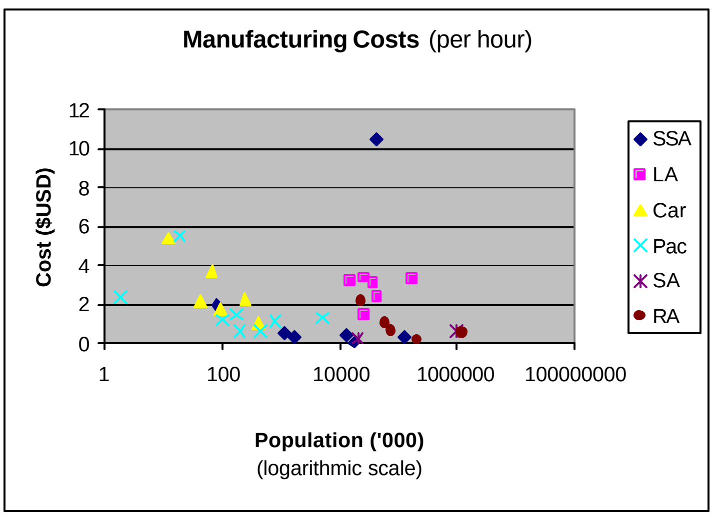
The following graph makes this point very clear, by plotting manufacturing costs against population, highlighting the OECD countries and the 3 high-income Asian entries.

Figure 4.12 Manufacturing Costs vs. Population



Thus, we need to eliminate these observations (from our already reduced dataset), similarly to what we have done for wages. The plot below shows the remainder of the observations by region. South Africa is the outlier.

Figure 4.13 Manufacturing Costs vs. Population (Small Sample)



There is perhaps a general declining trend in manufacturing labour costs as countries get larger, but with the (serious) exception of Latin America and South Africa where we seem to have higher manufacturing costs compared to the other countries in the same population categories.

Analysis of Size

In line with the treatment of the other cardinal variables, we have used the following specification to explore the influence of size:

$$\text{Ln(Manufacturing Costs)} = \alpha_0 + \alpha_1\text{*Ln(GDPc)} + \alpha_2\text{*Ln(Pop)} + \alpha_3\text{*[Ln(Pop)]}^2$$

Alternatively, we exclude the square term.

Results

Table 4.35 gives the results with all the observations available,

Table 4.35 Results for Manufacturing Costs (full sample)

	Manufact.		Manufact.
constant	**-0.503**	**constant**	**-0.685**
	-1.458		*-3.378*
LnGDPc	**0.963**	**LnGDPc**	**0.957**
	22.102		*22.583*
LnPop	**-0.030**	**LnPop**	**0.029**
	-0.324		*1.425*
LnPop2	**0.004**	-	
	0.651		
F-test (pop)	1.2173[.304]	-	
R-squared	0.90	R-squared	0.90
obs.	61	obs.	61

South Africa (+) and Botswana (-) were eliminated due to normality problems.

and table 4.36 excluding the OECD countries, Hong Kong, Singapore and Taiwan. Again, there are no significant size effects.

Table 4.36 Results for Manufacturing Costs
(Sub-sample excluding OECD and high-income Asia)

	Manufact.		Manufact.
constant	**-0.189**	**constant**	**-0.584**
	-0.317		*-1.538*
LnGDPc	**0.851**	**LnGDPc**	**0.853**
	6.533		*6.580*
LnPop	**-0.103**	**LnPop**	**0.025**
	-0.674		*0.652*
LnPop2	**0.008**	-	
	0.865		
F-test (pop)	.58530[.563]	-	
R-squared	0.64	R-squared	0.63
obs.	34	obs.	34

Conclusion

There is no persuasive evidence here of a small country cost disadvantage so we do not make any further calculations. The question is an important one, but with so few observations and such a vague question we are not sure that future work will yield much fruit. Certainly we are not inclined to give the absence of a size effect here precedence over our earlier results suggesting cost disadvantages to smallness. Unfortunately, we do not have productivity data to check whether the issue is just that higher wages in small countries are off-set by higher productivity.

IV. The Costs of Smallness
4. Utilities

The Business Cost Surveys ask utilities to provide up to four sorts of information: sunk (one-off), fixed costs, marginal costs and qualitative data on reliability. We analyse all four when they are available, but cannot effectively combine them into a single measure of utility costs. We find that the various elements are generally uncorrelated and that the data are both noisier and genuinely more variable than those analysed so far.

We also noted that the Survey does not include a question for water and electricity monthly charges. Therefore, as opposed to telephone where we have line rental fees, there is an important element missing when we analyse these services.

4.1 - Telephone

Data Sources

Data on GDP and Population were obtained from the same sources as previously. Information on telephone costs was taken from the Survey. Details on air distances are given in Annex IV.1 above.

Data concerns

There were several cases where countries charge, and hence report, a fixed fee per local call instead of a price per minute. These cases had to be excluded in order to avoid biasing the per-minute results. Furthermore, there are five countries that stated that local phone calls were free. We will see the impact of excluding these below. The USA reports call charges of 9 cents per minute for local calls. Since very local calls are free, we infer that this refers to national calls. This in turn raises a conceptual concern – how local is 'local'? Very local for very small countries but not necessarily for larger ones. Thus our size effects could include an element of internal distance.

Table 4.37 Missing Observations for Local Telephone Calls

	Local Calls	Comment from Survey

Missing obs.	Barbados	There is a flat rate charged for local calls, therefore calls are not charged by the minute.
	Nauru	Limited telephone service. Telephone service is not available in Nauru due to cable fault and non-payment of accounts to Australian Telecom.
	Canada	No per minute charge for local calls.
	Australia	Local calls on business accounts are untimed.
	Anguilla	Local calls are charged per call and not by the minute
	Kiribati	Flat rate per call.
Zeros	Hong-Kong	Local calls in Hong Kong are free.
	Marshall Is.	Nil.
	Palau	0.
	Philippines	Free in Manila.
	Micronesia	No rate for local calls - fixed monthly fee covers all local charges.

Regarding international calls, there are no observations for Nauru (for any of the three destinations) nor for calls from Cameroon to Tokyo. In addition, London, Tokyo and US do not figure in their respective regressions.

Table 4.38 Missing Observations for Fixed Costs

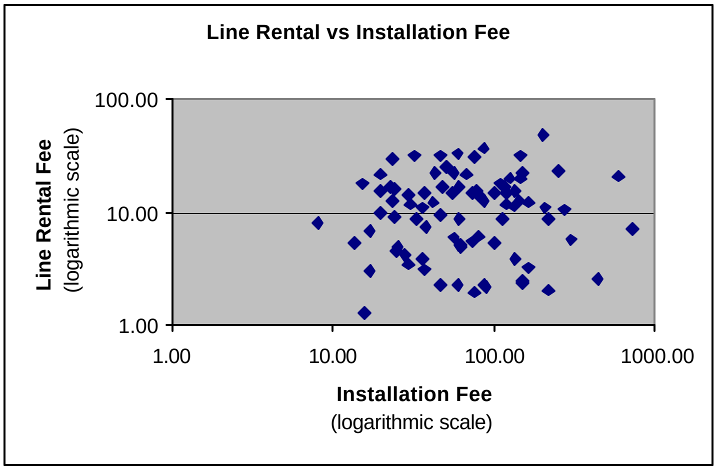
	Installation Fee	Line Rental
Missing obs.	Chile	Nigeria South Korea
Zeros	Thailand	Turkey US

Descriptives

We have plenty information regarding telephone costs. There are costs for international calls (Tokyo, London and New York), local calls, line rental fees and installation fees. Additionally we also have questions on disruptions to telephone services, repair and time to get a new telephone line (these are categorical variables and will be treated in the same way as the availability of workers, i.e., analysed through cros s-tabulations).

As previously, we first plotted the data for each variable. We then also plotted line rental fees against installation fees, installation fees against cost of local calls, and the cost of local calls against line rental fees. This procedure was designed to reveal any relationships between these variables, (e.g., a negative relationship between line rental and installation fees), but we found none.

Figure 4.14 Line Rental vs. Installation Fee



We also correlated the variables in alternate pairs to check this result. There is no apparent relationship between the sunk, fixed and marginal costs of telephone services in these data. This makes overall cost comparisons very difficult and we attempt none.

Figure 4.15 Local Calls vs. Line Rental Fee

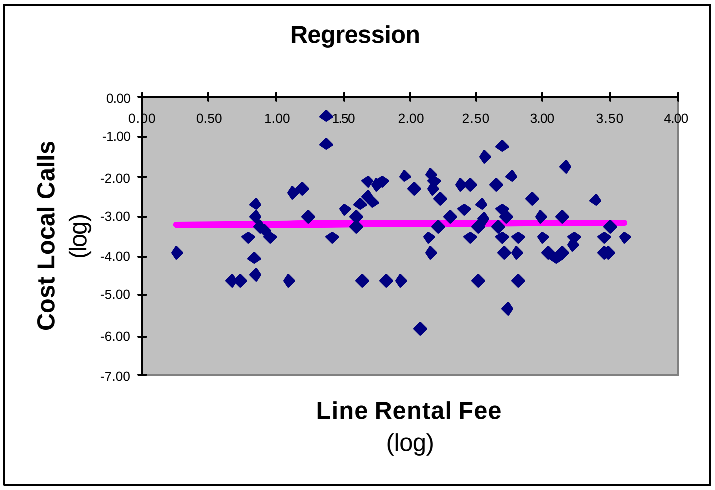
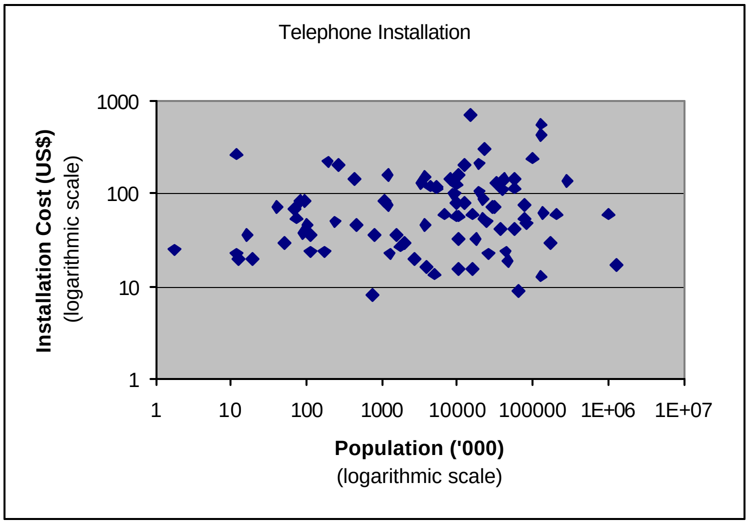
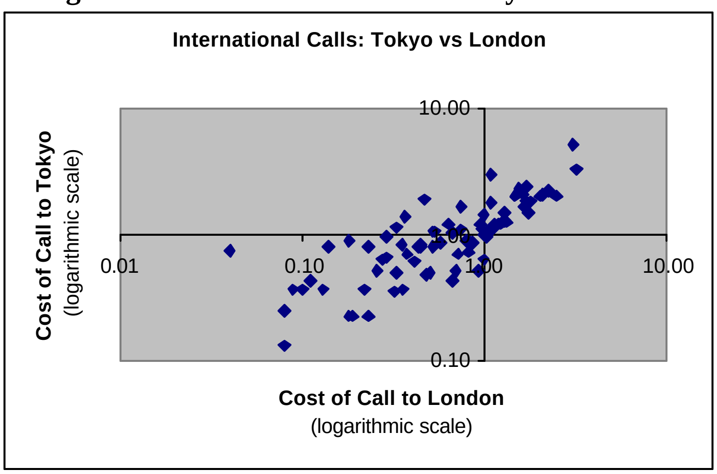


Figure 4.16 Installation Fee vs. Population



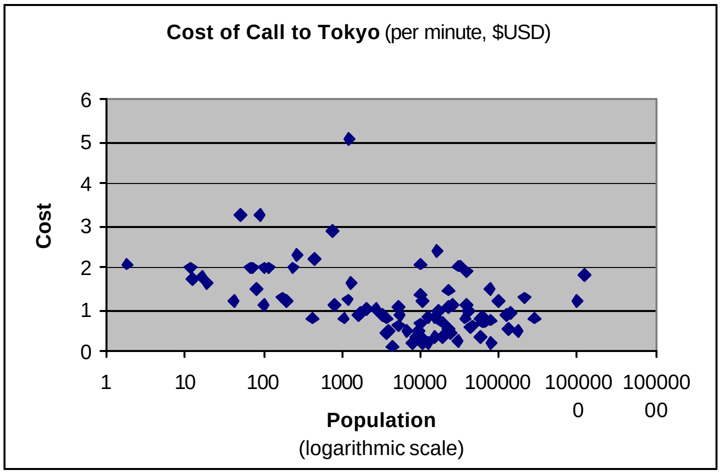
We also have the cost of international calls to London, Tokyo and New York. We plotted each pair combination to observe the respective correlations. Given the dispersion of the three destinations, the positive correlations suggest that there are factors other than geographical distance to the destination city that influence international call prices. We interpret these as country effects specific to our target countries.

Figure 4.17 International Calls: Tokyo vs. London



Finally, we plotted costs of international calls against population. There is a discernable downward slope, indicating that population may have a significant negative impact on the costs of international calls.

Figure 4.18 Cost of Calls to Tokyo vs. Population



Analysis of Size

As was done before, the variables were logged for these regressions. The sample of each regression varies since there are different missing values for different telephone costs. We have data on both local and international telephone costs, and use different specifications to study the effect of population on these costs.

For international telephone costs we have used the same specification as in the airfreight analysis.

$$\text{Ln(International Telephone Costs)} = \alpha_0 + \alpha_1\text{*Ln(Air Distance)} + \alpha_2\text{*Ln(Pop)} + \alpha_3\text{*[Ln(Pop)]}^2$$

For local telephone costs and fixed costs, we have used GDP per capita and Population as regressors:

$$\text{Ln(Costs)} = \alpha_0 + \alpha_1\text{*Ln(GDPc)} + \alpha_2\text{*Ln(Pop)} + \alpha_3\text{*[Ln(Pop)]}^2$$

Results

The results for international calls are reproduced in table 4.39:

Table 4.39 Results for International Telephone Costs

	London[1]	Tokyo	NY
constant	**-1.777**	**1.196**	**-0.676**
	-1.844	*1.132*	*-0.590*
LnDistance	**0.380**	**0.076**	**0.155**
	4.645	*0.647*	*1.322*
LnGDPc	**-0.231**	**-0.185**	**-0.383**
	-4.599	*-4.497*	*-8.032*
LnPop	**-0.338**	**-0.345**	**-0.394**
	-3.089	*-3.341*	*-3.592*

LnPop2	**0.015**	**0.015**	**0.016**
	2.215	*2.256*	*2.261*
F-test on Pop	11.65 [.000]	17.81 [.000]	23.96 [.000]
R-squared	0.63	0.40	0.61
obs.	89	89	90

[1] New Zealand (-) was excluded due to normality problems
All these regressions fail Functional Form

From table 4.39 we see that Population has a significant negative effect on the cost of calls. The F-tests on Population prove the existence of a quadratic relationship between population and costs. However the Functional Form test fails for all regressions. We tested other functional forms and observed that excluding the squared term in population cures the first two regressions. To cure the problems for New York we would have to exclude distance and population squared simultaneously. However, and since Pop2 is statistically significant, we have decided to keep it, although do below compute the cost disadvantages from the regressions without Pop2 as a comparison.

Table 4.40 Results for Local Telephone and Fixed Costs

	Local[1]	**Installation Fee**	**Line Rental**
constant	**-2.471**	**3.524**	**2.039**
	-6.451	*11.385*	*9.119*
LnGDPc	**-0.074**	**0.106**	**0.364**
	-1.025	*1.633*	*7.654*
LnPop	**-0.073**	**0.058**	**-0.026**
	-1.800	*1.721*	*-1.057*
R-squared	0.05	0.06	0.42
obs.	81	90	88

[1] Five countries stated '0' cost for local calls. These were excluded since it is not possible to log them.

Regarding the cost of local telephone calls, GDP per capita seems not to be significant but we leave it in place to avoid any omitted variables bias. The F-test indicates that, with the quadratic form, population is significant only at 10%, and since the squared term is not significant, we drop it. For fixed costs, we see that the population effect is significantly positive (at 10%) for installation fees but not for line rental fees, where GDPc seems to have an important role in explaining the prices.

<u>Sensitivity Tests</u>

There is a potentially serious bias in the equation above, because the five omitted observations with free calls are all small. If we replace their zeros by $0.003 per minute - the lowest positive value observed - there is no significant relationship at all. We would conclude, therefore, that size is not a determinant of local telephone costs.

<u>Cost of Smallness</u>

We proceed to calculate the cost disadvantages for telephone costs in table 4.41.

Table 4.41 Cost Deviation from the Median Country (%) for Telephone Costs

Size	Pop ('000)	London	Tokyo	NY	Local	Install. Fee	Line Rental
Micro	12.13	197.6	211.9	300.7	*0*	-32.3	*19.1*
Very Small	197.00	60.6	65.1	89.1	*0*	-20.4	*10.8*
Small	4,018.00	7.1	7.8	11.0	*0*	-5.2	*2.4*
T point		78,171.99	98,715.77	222,459.4			
Closest Country		Vietnam	Mexico	Indonesia			

It is clear that the cost disadvantages facing the small countries for international calls are quite heavy. The parallels with airfreight are strong and these results reinforce the view that isolation is a major problem for small economies. On the other hand installation seems cheaper in smaller countries.

Table 4.42 Cost Deviation from the Median Country (%) for International Calls with the Log-linear equation

Country	Pop ('000)	London	Tokyo	NY
Micro	12.13	97.1	119.4	177.6
Very Small	197.00	48.7	58.4	81.7
Small	4,018.00	9.7	11.3	14.9

As we would expect, the major differences between the linear and quadratic forms are in the extremes of the distribution, with the latter implying much higher cost disadvantages for very small countries. However, for Vanuatu the difference is not that big, while for Singapore the cost disadvantage is greater for the (log-)linear variant.

Analysis of the qualitative data

We have qualitative information in the Survey, on disruptions in telephone services, time to repair faults and the availability of new telephone lines.

Table 4.43 Data Decisions for the Categorical Variables (Telephone)

	Country	Question	Comment	Decision
Missing obs.				
Others	Kenya	New Connection	(Stated 'No' but also 'More than a month')	No
	Nigeria	New Connection	(Stated 'Not always' but also 'More than a month')	No
	Suriname	New Connection	(Stated 'No' but also 'More than a month')	No
	Suriname	Repair	5 working days and two weeks.	2 weeks

As before, we cross-classified observations with Population and GDP.

The Wait for New Connections

The Business cost survey asked respondents to report the wait for a new line in five categories:

(1) - <72 hours
(2) - < 1 week
(3) - 1 week - 1 month
(4) - >1 month
(5) - Not available (i.e., indefinite delays)

Table 4.44 Cross Tabulations for New Telephone Connections (Population)

Wait for New Connection	Population					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
<72 hours	2	3	6	9	2	22
<1 week	5	3	3	7	7	25
1 week - 1 month	11	3	4	11	6	35
>1 month	2	-	1	2	1	6
Not Available	-	1	-	1	2	4
Total	20	10	14	30	18	92

Table 4.45 Cross Tabulations for New Telephone Connections (GDP)

Wait for New Connection	GDP					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
<72 hours	1	2	3	4	12	22
<1 week	4	2	2	4	13	25
1 week - 1 month	9	4	9	6	7	35
>1 month	2	1	2	1	-	6
Not Available	-	1	-	3	-	4
Total	16	10	16	18	32	92

In 35% of the small economies users can have a new line within a week. This is by far the lowest percentage compared with the other population categories. The bulk of very small country observations (55%) is agglomerated in category 3 (1 week to 1 month). There are clearly some business activities for which this is a serious problem, but others for which it is probably not.

As before, we group the data into two 2-by-5 tables to test the significance of the apparent associations.

Table 4.46 Chi-Square Statistics for New Telephone Connections

Type of Grouping	Pearson's Chi Square		
	Value	df	Asymp. Sig. (2 sided)
Pop (1, 2-5)	4.368	4	0.359
Pop (1-2, 3-5)	3.434	4	0.488
GDP (1, 2-5)	10.245	4	0.036
GDP (1-2, 3-5)	15.211	4	0.004

These tests indicate that aggregate GDP does appear to be a significant factor in connection times whilst the same cannot be said about population. Taking the latter more seriously for consistency's sake, we conclude that there is no difference in the delay to install a new line in the small countries and larger ones. This result is reinforced with the non-significance of population in the ordered logit.

Disruption

Respondents classified disruptions by five categories:

(1) - No disruption
(2) - Very rare
(3) - Rare
(4) - Quite frequently
(5) - Most frequently

Table 4.47 Cross Tabulations for Telephone Disruptions (Population)

Frequency of Disruptions	Population					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
No disruption	4	-	6	11	8	29
Very rare	12	6	6	14	8	46
Rare	3	2	2	4	1	12
Quite Frequently	-	1	-	1	1	3
Most Frequently	1	1	-	-	-	2
Total	20	10	14	30	18	92

Table 4.48 Cross Tabulations for Telephone Disruptions (GDP)

Frequency of Disruptions	GDP					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
No disruption	3	-	3	6	17	29
Very rare	8	7	8	9	14	46
Rare	4	1	5	1	1	12
Quite Frequently	-	1	-	2	-	3
Most Frequently	1	1	-	-	-	2
Total	16	10	16	18	32	92

There are hints of a size effect here. Small countries have lower percentages of 'no disruption', while having higher percentages (15%) for rare and quite frequent disruption. In general, there are few instances of very severe disruption of telephone services that can affect the reliability of business: five cases in all.

Collapsing the data into 2-by-5 tables and testing for association produces mixed results.

Table 4.49 Chi-Square Statistics for Telephone Disruptions

Type of Grouping	Pearson's Chi Square		
	Value	df	Asymp. Sig. (2 sided)
Pop (1, 2-5)	8.427	4	0.077
Pop (1-2, 3-5)	3.983	4	0.408
GDP (1, 2-5)	13.967	4	0.007
GDP (1-2, 3-5)	9.393	4	0.052

Again GDP seems to be much more significant correlate of disruption than Population. However, we can see some evidence of higher disruptions (at 10%) if we analyse ‘no disruption’ against the other four categories.

The results of table 4.50 reinforce this conclusion. The ordered logit model generated significant population effects and the predictions show that countries with population below 9 million inhabitants have greater probability of facing disruptions in the telephone service.

Table 4.50 Ordered Logit for Telephone Disruptions

Telephone Disruptions	Population (‘000) [if GDP pc = $10,000]
1	Above 9,170
2	Below 9,170
3	
4	
5	

Repair

Respondents classified times for repair into four categories:

(1) - 48 hours
(2) - 5 working days
(3) - 2 weeks
(4) - > 2 weeks

Table 4.51 Cross Tabulations for Telephone Repair Times (Population)

Repair Times	Population					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
48 hours	13	8	10	21	14	66
5 working days	5	-	2	5	2	14
2 weeks	1	2	1	3	1	8
>2 weeks	1	-	1	1	1	4
Total	20	10	14	30	18	92

Table 4.52 Cross Tabulations for Telephone Repair Times (GDP)

Repair Times	GDP					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
48 hours	10	7	7	13	29	66
5 working days	4	1	5	2	2	14
2 weeks	1	2	2	2	1	8
>2 weeks	1	-	2	1	-	4
Total	16	10	16	18	32	92

Only 65% of the small states provide repairs within 48 hours, which is the lowest percentage amongst population groups, but only just. Overall there is little apparent evidence that repairs take longer in small countries, although the formal tests again suggest a relationship with GDP for repairs within 48 hours.

Table 4.53 Chi-Square Statistics for Telephone Repair Times

Type of Grouping	Pearson's Chi Square		
	Value	df	Asymp. Sig. (2 sided)
Pop (1, 2-4)	1.154	4	0.886
Pop (1-2, 3-4)	0.671	4	0.955
GDP (1, 2-4)	12.054	4	0.014
GDP (1-2, 3-4)	5.431	4	0.246

Overall the tables and tests suggest that there is no significant evidence of excess delays in repair times in small countries. Again, the ordered logit confirms this conclusion.

Conclusion

We conclude that there is evidence of a strong population effect on the cost of international calls. We also have evidence of higher disruptions in telecommunication services in the small states represented in the dataset – especially when measuring size by GDP. There is not, however, much evidence that the quality of the telephone service or the costs of local calls vary by size.

IV. The Costs of Smallness
4. Utilities
4.2 - Electricity

Data Sources

Data on GDP and Population was obtained from the sources mentioned above. Information on electricity costs was taken from the Survey.

Data concerns

We classified four observations as 'missing' and there was one explicit zero in the sample (Nauru). These countries will not enter our analysis. Some figures are the simple average of different rates given for different levels of usage.

Table 4.54 Missing Values for Electricity Costs (per kilowatt)

	Electricity	Comment

Missing obs.	Gabon	Electricity charges are paid in advance of use, varying from CFAfr220,000 to CFAfr430,000.
	Hong Kong	The cost of electricity (commercial tariff) is HK$99.60 per month for the first 1,500 units and HK$107.60 per month for usage above this level.
	Senegal	The amount paid for electricity varies depending on degree of usage.
	Seychelles	US$207.00/KVA average.
Zeros	Nauru	Free.

Here are some of the decisions we had to make regarding electricity connection fees.

Table 4.55 Data Decisions for Electricity Costs (per kilowatt)

Country	Remarks	Decision
Antigua & B.	US $5.56 installation; US $3.70 for inspection; cost of wire (0.24/foot).	$9.26
Denmark	600/ampere	N/A
Indonesia	0.03 - This figure is per VA of capacity	N/A
Singapore	There is no electricity reconnection fee for an established property. Connection fee for new developments (supply directly from grid) is S$1600.	0
South Korea	$57.61 (Connenction charge is for 1km of overhead connection. Price rises to W101,200 for 1km underground connection).	0
Sweden	There is no connection fee for electricity, but an annual fee is charged, depending on usage.	0
UK	Connection fee is incurred as part of a standing charge of 14.08 pence per day	N/A

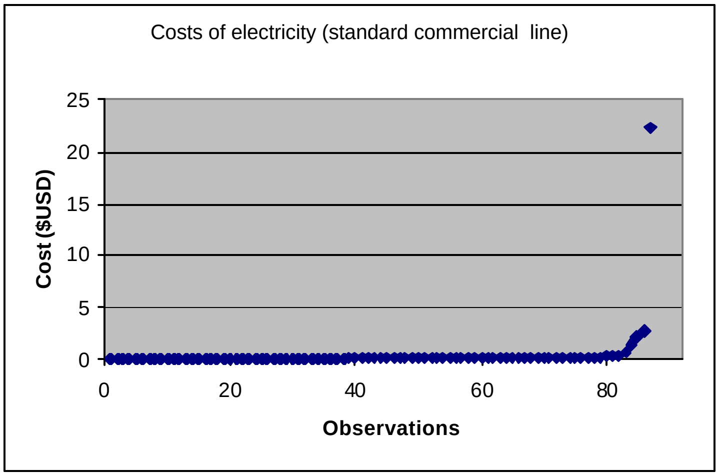
Table 4.56 Missing Observations for Electricity Connection Fees

	Electricity Connection Fees	
Missing obs.	Belgium	Ireland
	Cameroon	Italy
	Denmark	Malaysia
	Germany	St Kitts & N.
	Grenada	Trinidad & T.
	Hong Kong	UK
	India	Venezuela
	Indonesia	Zimbabwe
Zeros	Australia	Senegal
	Chile	Singapore
	China	South Korea
	Finland	Sweden
	Nauru	

Descriptives

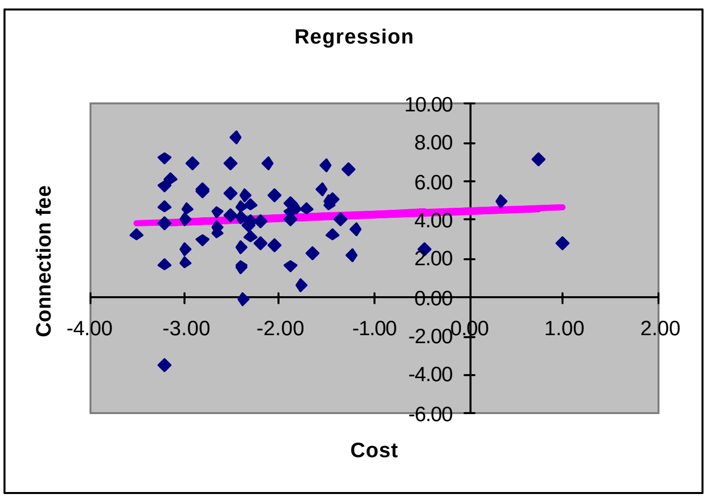
The graph below shows the plot of the marginal cost of electricity. Poland, with US$22.38 per kilowatt of electricity, is a clear outlier. This value is most probably expressed in cents, so we decided to divide it by 100, which brings it right back into line.

Figure 4.19 Ele ctricity Costs (per kilowatt)



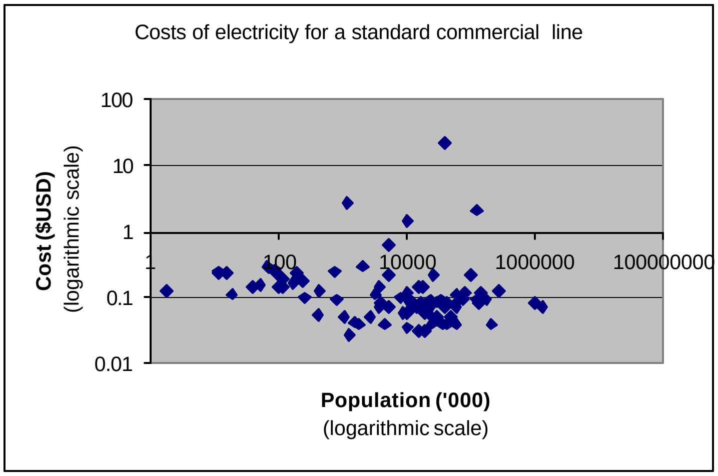
We then correlated the costs of electricity and the connection fees. As with the telephones, one might find reasons to expect either a negative or a positive relationship, but, in fact, they seem not to be related.

Figure 4.20 Connection Fee vs. Electricity Costs (per kilowatt)



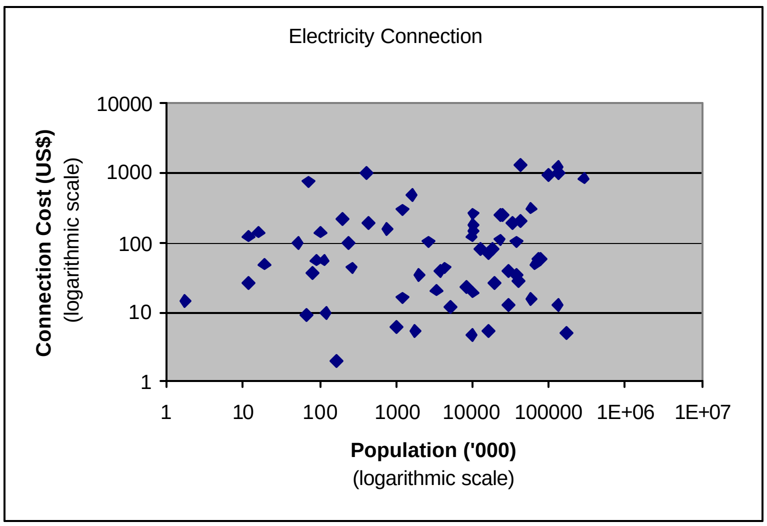
Below we present the plot that shows the possible relationship between electricity costs and population. We can see that, ignoring the outliers, there might be a slightly negative trend.

Figure 4.21 Electricity Costs (per kilowatt) vs. Population



The following plot, however, shows no clear correlation between electricity connection fees and population, however.

Figure 4.22 Connection Fee vs. Population



Analysis of Size

For electricity costs we have used the following standard specifications:

$$Ln(\text{Electricity Costs}) = \alpha_0 + \alpha_1 * Ln(GDPc) + \alpha_2 * Ln(Pop) + \alpha_3 * [Ln(Pop)]^2$$

$$Ln(\text{Electricity Costs}) = \alpha_0 + \alpha_1 * Ln(GDPc) + \alpha_2 * Ln(Pop)$$

Results

Nauru was deleted since it had '0' for costs of electricity, which could not be logged. Nevertheless, we carried sensitivity tests to assess the implications of this procedure.

We report the results in table 4.57 without the squared term, since it turned out to be not significant.

Table 4.57 Results for Electricity Costs

	Electricity[1]	Connection Fee
constant	**-1.479**	**3.536**
	-7.105	*5.826*
LnGDPc	**-0.021**	**-0.223**
	-0.493	*-1.589*
LnPop	**-0.098**	**0.103**
	-4.350	*1.519*
R-squared	0.19	0.07
obs.	84	67

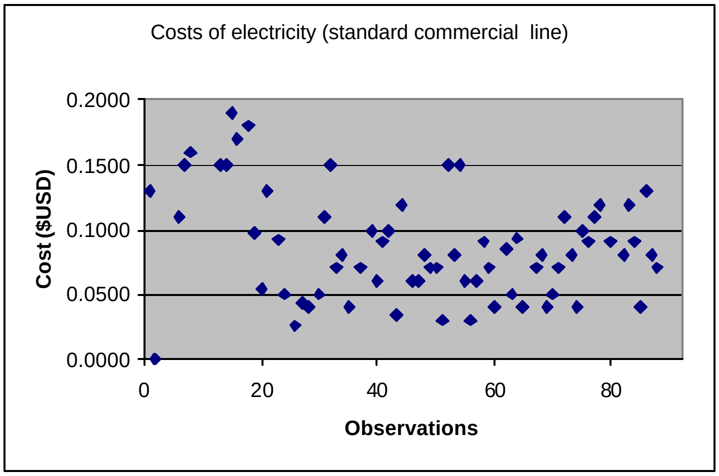
[1] Mauritius (+), Nigeria (+) and Czech Republic (+) were also eliminated since they caused normality problems.

Population seems to have a strong negative effect on the marginal cost of electricity. However, GDP per capita seems not as useful explaining different prices per kilowatt. On connection fees we have (weakly and noisily) the opposite story: a decline in cost with GDP pc and an increase with population, although neither co-efficient is significant at 5%.

Sensitivity Tests

In this section we tried to assess the importance of dropping the zero (Nauru). This country is shown below, together with the lower part of the cost sub-sample.

Figure 4.23 Electricity Costs (per kilowatt) (Lower Sub-Sample)



If we set Nauru at US$0.01 (which is around a third of the value for smallest non-zero observation) and proceed as usual with eliminating outlying observations, we lose Mauritius, Nigeria and Czech Republic as before. The coefficient on population at this stage has fallen (absolutely) from -0.098 to -0.074, but Nauru is a plain outlier.

The main conclusion that we draw from these tests is that the coefficient on population is still significantly negative. Thus, although we should be concerned about excluding Nauru, so far as the precise magnitude is concerned, the broad thrust of the results is unchanged.

Cost of Smallness

We have used the same procedure as before for calculating the cost disadvantage of the countries. Below (table 4.58) we have calculated how much that values deviates from the median country (observation 46, which has a population of approximately 10 million inhabitants)

Table 4.58 Cost Deviation from the Median Country (%) for Electricity Costs

Size	Pop ('000)	Electricity	Connection Fee
Micro	12.13	93.1	*-49.9*
Very Small	197.00	47.0	*-33.3*
Small	4,018.00	9.4	*-9.0*

Again the cost disadvantage on the marginal cost of electricity is very large for micro economies, although it is partly offset by (insignificantly) lower connection fees. As noted above, however, we do not have the information necessary to make a precise trade-off here.

Analysis of the qualitative data

We now analyse the qualitative information in the Survey, regarding disruptions in power supply and the availability of new electricity connections. As always, there were some decisions we had to make concerning the data.

Table 4.59 Data decisions for the Categorical Variables (Electricity)

	Country	Question	Comment	Decision
Missing obs.	India	New Connection		
Others	Vanuatu	Disruptions	Rare to very rare. Sometimes due to maintenance of network, more frequently for bills' outstanding.	Rare
	Guyana	Disruptions	Rare (in some areas) and quite frequently (in some areas)	Quite Freq.
	Trinidad & T.	Disruptions	Very rare (applies to most areas) and rare (applies to some areas)	Very rare

Disruption

Respondents classified disruptions by five categories:

(1) - No disruption
(2) - Very rare
(3) - Rare
(4) - Quite frequently

(5) - Most frequently

Table 4.60 Cross Tabulations for Electricity Disruptions (Population)

Frequency of Disruptions	Population					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
No disruption	-	1	5	11	5	22
Very rare	9	4	5	10	5	33
Rare	6	3	2	3	3	17
Quite Frequently	4	1	2	6	4	17
Most Frequently	1	1	-	-	1	3
Total	20	10	14	30	18	92

Table 4.61 Cross Tabulations for Electricity Disruptions (GDP)

Frequency of Disruptions	GDP					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
No disruption	-	-	2	5	15	22
Very rare	4	5	5	6	13	33
Rare	6	3	4	2	2	17
Quite Frequently	4	2	5	4	2	17
Most Frequently	2	-	-	1	-	3
Total	16	10	16	18	32	92

With these tables we observe a clear tendency for small countries to have relatively more disruptions in electrical supply: 25% of them have frequent power disruption (although the same is true of population category 5) but none has completely disruption-free services. Testing formally as before:

Table 4.62 Chi-Square Statistics for Electricity Disruptions

Type of Grouping	Pearson's Chi Square		
	Value	df	Asymp. Sig. (2 sided)
Pop (1, 2-5)	11.251	4	0.024
Pop (1-2, 3-5)	4.442	4	0.349
GDP (1, 2-5)	18.738	4	0.001
GDP (1-2, 3-5)	20.398	4	0.000

These tests show that small countries do appear to suffer more disruption problems than more populated countries. Table 4.63, which reports the implications of a significant population co-efficient in the ordered logit, also suggest strong disruptions increase with lower levels of population. The effect can be severe for countries with less than 1.4 thousand inhabitants.

Table 4.63 Ordered Logit for Electricity Disruptions

Electricity Disruptions	Population ('000) [if GDP pc = $10,000]
1	Above 58,832
2	From 1.4 to 58,832

3	Below 1.4
4	
5	

The Wait for New Connections

Respondents reported the wait for a new line in five categories:

(1) - <72 hours
(2) - < 1 week
(3) - 1 week - 1 month
(4) - >1 month
(5) - Not available (i.e., indefinite delays)

Table 4.64 Cross Tabulations for New Electricity Connections (Population)

Wait for New Connection	Population					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
<72 hours	4	2	7	8	3	24
<1 week	1	3	3	5	4	16
1 week - 1 month	13	4	4	10	6	37
>1 month	2	1	-	5	4	12
Not Available	-	-	-	2	-	2
Total	20	10	14	30	17	91

Table 4.65 Cross Tabulations for New Electricity Connections (GDP)

Wait for New Connection	GDP					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
<72 hours	4	-	5	2	2	24
<1 week	1	3	-	6	6	16
1 week - 1 month	10	7	5	6	6	37
>1 month	1	-	5	3	3	12
Not Available	-	-	1	1	1	2
Total	16	10	16	18	18	91

Table 4.84 and 4.85 show that 65% of the very small states have to wait between 1 week and 1 month to obtain a new connection. Notwithstanding this, however, there is no compelling evidence of differences in delays in the supply of new connections. Neither the Chi-squared tests nor the ordered logit model uncover significant differences across countries.

Table 4.66 Chi-Square Statistics for New Electricity Connections

Type of Grouping	Pearson's Chi Square		
	Value	df	Asymp. Sig. (2 sided)
Pop (1, 2-5)	4.158	4	0.385
Pop (1-2, 3-5)	7.413	4	0.116
GDP (1, 2-5)	3.447	4	0.486
GDP (1-2, 3-5)	6.671	4	0.154

Conclusion

Throughout this section it has been clear that small states face a heavy penalty in terms of marginal electricity costs, and are also probably subject to reduced reliability in the service regarding disruptions. Given the economies of scale inherent in electrical generation, this is not surprising. There appear to be no significant extra delays for connection in small countries, however, and only weak evidence that they face lower connection fees.

IV. The Costs of Smallness
4. Utilities
4.3 - Water

Data Sources

Data on GDP and Population was obtained from the same sources as before. Information on the costs of water was taken from the Survey.

Data concerns

Table 4.67 reports the main difficulties found with this dataset. We classified four countries as N/A, whereas another four presented '0' for costs of water.

Table 4.67 Missing Observations for Water Usage Costs

	Water	Comment
Missing obs.	Argentina	There is no fee for water consumption per litre. Companies pay a bimonthly tax depending on size of property. A small office (180 sq metres) would pay approximately Ps40 bimonthly.
	Gabon	Water is charged for monthly.
	Pakistan	Companies using water for drinking and toilets only are charged PRs600 per month. Industries where water is used in manufacturing are charged PRs1200 per month. Industries where water is used as an input are charged PRs3000 per month.
	Japan	$145.2 per month.
Zeros	Cook Is.	Free at present.
	Germany	There is no charge for water supply; though a charge for wastewater is included in industrial rent.
	Niue	Free. Water supply is from rain and desalination. In the event of a draught and failure of power supply the desalination plant will not function thus no water will be available. If there is no rain and the desalination plant malfunctions then there will be no electricity, as the power cooling system will not function. All these are intertwined to make life in Nauru so difficult.
	Nauru	Free.

Table 4.68 Data Decisions for Water Usage Costs

Country	Remarks	Decision
South Korea	1.47US$/cubic metre	N/A
Trinidad & T.	US$ 2000 - US$ 3000 (metered) – 100 mm diameter connection	$2500
US	$1290 (The connection fee given is for a I inch pipe).	$1290

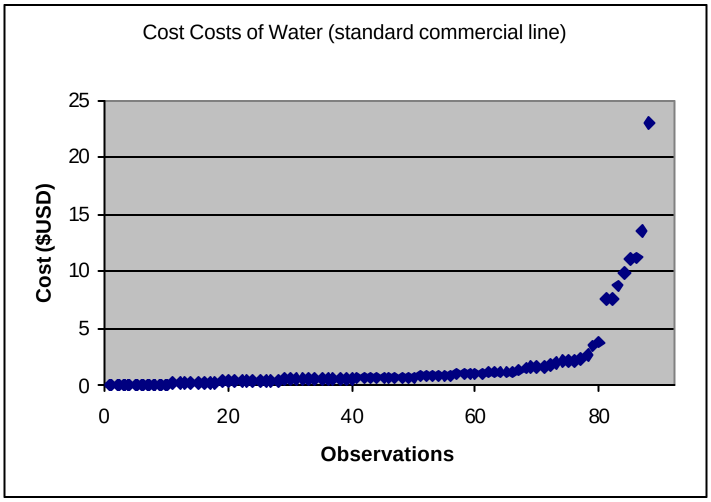
Table 4.69 Missing Observations for Water Connection Fees

	Water Connection Fees	
Missing obs.	Argentina	Nigeria
	Brazil	Poland
	Cameroon	South Korea
	France	St Kitts & N.
	Hong Kong	St Vincent & G.
	India	Tuvalu
	Malaysia	Zimbabwe
	Nauru	
Zeros	Australia	Germany
	Canada	Japan
	Chile	Namibia
	China	Singapore
	Czech Repub.	Sweden
	Denmark	

Descriptives

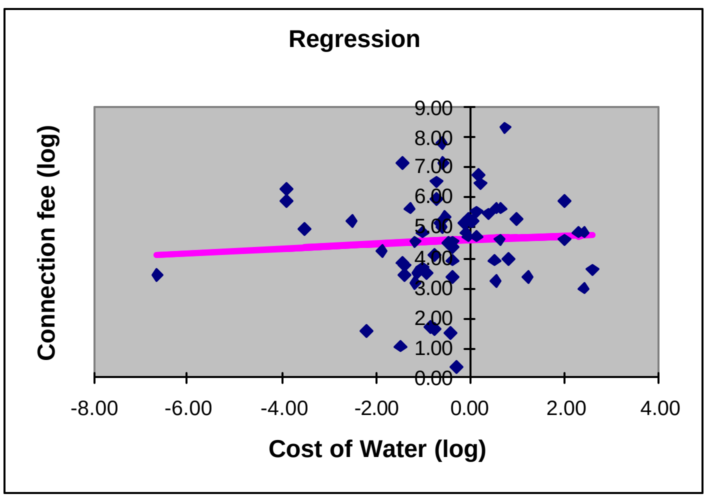
The graph below suggests that we may have to deal with some outliers (Tuvalu is the highest with US$23.00), but on investigation we found no evident typing errors. Hence we carried the extreme values forward to the next stage of the analysis. If the model does not explain these high values, they will very probably violate the normality assumption and will be subsequently deleted

Figure 4.24 Costs of Water (per 1000 litres)



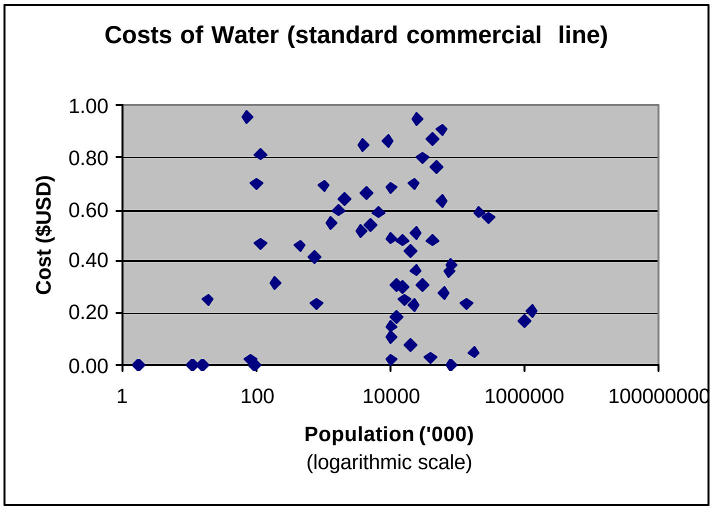
As before we regressed cost of water on water connection fees, and again as before there is no apparent relationship between these two variables (i.e. the fixed and variable costs).

Figure 4.25 Connection Fee vs. Costs of Water (per 1000 litres)



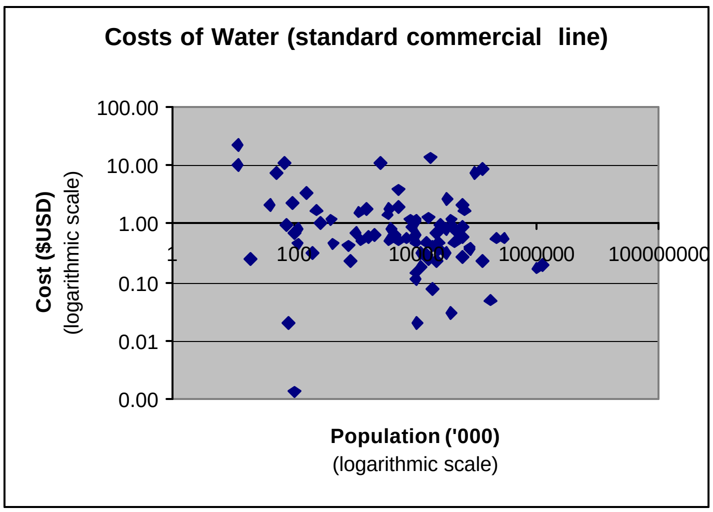
Below we plot the marginal cost of water (per 1000 litres) against population for the lower-cost sub-sample. This procedure is intended to show how far the zeros are from the rest of the observations.

Figure 4.25 Costs of Water (per 1000 litres) vs. Population (Lower Sub-Sample)



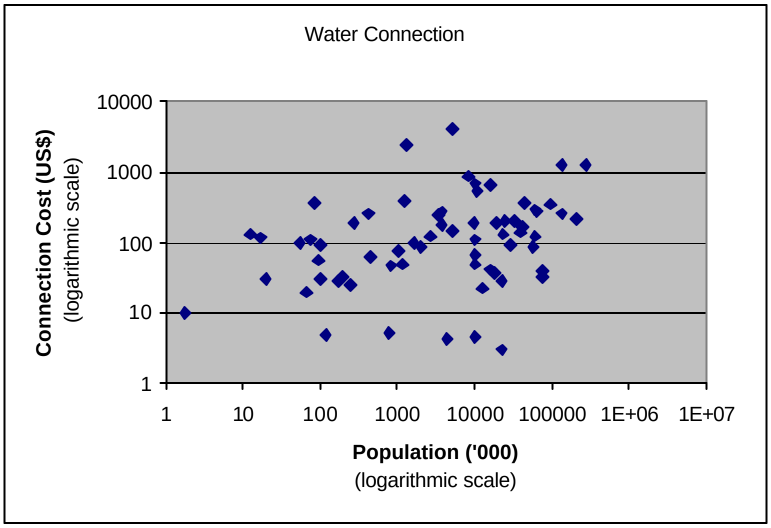
Finally, we look for the population effect on the costs of water. The plot below hints at a negative relationship between these two variables, but, critically, the figure excludes the zeros because they cannot be logged.

Figure 4.26 Costs of Water (per 1000 litres) vs. Population



The corresponding plot for connection fees suggests a positive relationship – i.e. advantages to smallness – as hinted at for electricity.

Figure 4.27 Connection Costs vs. Population



Analysis of Size

For the costs of water we have used the following specification:

$$\text{Ln(Water Costs)} = \alpha_0 + \alpha_1 \text{*Ln(GDPc)} + \alpha_2 \text{*Ln(Pop)} + \alpha_3 \text{*[Ln(Pop)]}^2$$

and tested the squared term.

Results

Four countries were excluded since they had '0' for costs of water, but we return to them below. The squared term was not significant and was dropped. Note also that normality testing resulted in the omission of some further small and very low cost observations.

Table 4.70 Results for Water Costs

	Water[1]	**Connection Fee**[2]
constant	**1.286**	**2.645**
	2.649	*6.022*
LnGDPc	**0.024**	**0.409**
	0.260	*4.177*
LnPop	**-0.184**	**0.219**
	-3.516	*4.445*
R-squared	0.14	0.35
obs.	82	63

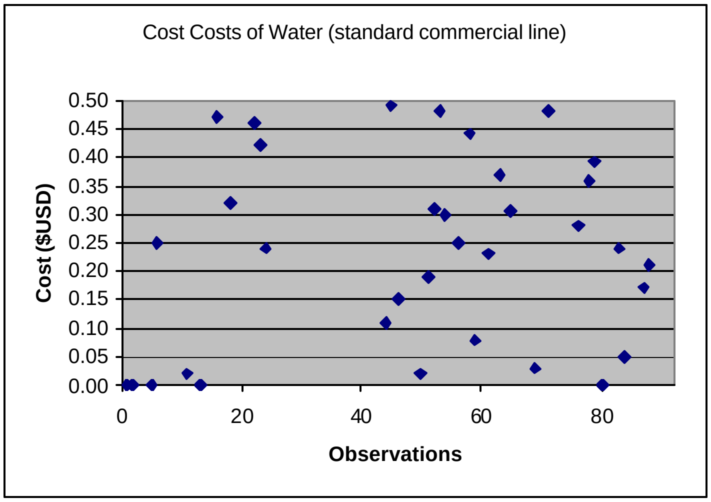
[1] Seychelles (-) and Grenada (-) were eliminated because they caused normality problems.
[2] Taiwan (-), Norway (-) and Portugal (-) were excluded due to normality problems.

As with telephone costs, GDP per capita is not significant. Nevertheless, there is evidence of a significant (and negative) effect of population on water costs.

Sensitivity Tests

As we have done before, we 'anchor' our zeros to the minimum value in the sample in order to evaluate the implications of their exclusion for the results presented above. Here the zeros were set to 0.0004, which is a third of the lowest observation (Grenada). Grenada stated US$0.0013 (plus US$2.93 service charge per month) and Seychelles US$0.02.

Figure 4.28 Connection Costs (Lower Sub-Sample)



The net effect of adding these observations back into the sample is to turn the population effect insignificantly positive, which seems to indicate that we cannot safely assume a relationship between water costs and size.

Cost of Smallness

There is a conflict in these results. The 'standard' equation shows very strong cost disadvantages to smallness effects (Table 4.71). However, they are subject to the caveat that six small low-cost countries were excluded, which almost certainly biases the results. Hence we conclude conservatively that there is no cost of smallness in water. For connection fees, on the other hand, there is quite strong evidence that fees increase with size.

Table 4.71 Cost Deviation from the Median Country (%) for Water Costs

Size	Pop ('000)	Water	Connection Fee
Micro	12.13	*244.1*	-77.0
Very Small	197.00	*106.1*	-57.7
Small	4,018.00	*18.3*	-18.1

Analysis of the qualitative data

Table 4.72 Data decisions for the Categorical Variables (Water)

	Country	Question	Comment	Decision
Missing obs.	South Korea India	Disruptions New Connection		
Others	Argentina	New Connection	(Stated 'Yes' for availability though not stating the duration).	N/A

Disruption

(1) - No disruption
(2) - Very rare
(3) - Rare
(4) - Quite frequently
(5) - Most frequently

Table 4.73 Cross Tabulations for Water Disruptions (Population)

Frequency of Disruptions	Population					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
No disruption	2	2	7	10	8	29
Very rare	11	5	2	12	5	35
Rare	1	2	5	4	3	15
Quite Frequently	4	-	-	3	2	9
Most Frequently	2	1	-	-	-	3
Total	20	10	14	29	18	91

Table 4.74 Cross Tabulations for Water Disruptions (GDP)

Frequency of Disruptions	GDP					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
No disruption	2	1	2	5	19	29
Very rare	7	5	7	8	8	35
Rare	-	4	6	3	2	15
Quite Frequently	4	-	1	2	2	9
Most Frequently	3	-	-	-	-	3
Total	16	10	16	18	31	91

Although the bulk of small economies (55%) face only very rare disruptions, only 10% admit not having disruption at all. Furthermore, 30% of very small countries have quite frequent disruptions (which can affect business running). The formal tests suggest some size effect on the probability of having no disruption (especially using GDP), but no effect on the distribution of observations about the rare/very rare boundary.

Table 4.75 Chi-Square Statistics for Water Disruptions

Type of Grouping	Pearson's Chi Square		
	Value	df	Asymp. Sig. (2 sided)
Pop (1, 2-5)	8.576	4	0.073
Pop (1-2, 3-5)	0.974	4	0.914

GDP (1, 2-5)	**20.229**	**4**	**0.000**
GDP (1-2, 3-5)	**7.759**	**4**	**0.101**

Complementing these tests, we have also run ordered logits. Contrary to the contingency tables, the results suggest that being small and having low levels of GDP significantly increases the probability of having water disruptions. Table 4.76 reports the exercise we have been using to illustrate the population impact on the predicted water disruption level.

Table 4.76 Ordered Logit for Water Disruptions

Water Disruptions	Population ('000) [if GDP pc = $10,000]
1	Above 3,206
2	From 0.4 to 3,206
3	Below 0.4
4	
5	

The Waiting Time New Connections

(1) - <72 hours
(2) - < 1 week
(3) - 1 week - 1 month
(4) - >1 month
(5) - Not available (i.e., indefinite delays)

Table 4.77 Cross Tabulations for New Water Connections (Population)

Wait for New Connection	Population					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
<72 hours	5	2	4	7	2	20
<1 week	4	4	4	4	6	22
1 week - 1 month	10	4	5	8	5	32
>1 month	1	-	1	7	3	12
Not Available	-	-	-	3	1	4
Total	20	10	14	29	17	90

Table 4.78 Cross Tabulations for New Water Connections (GDP)

Wait for New Connection	GDP					Total
	Cat 1	Cat 2	Cat 3	Cat 4	Cat 5	
<72 hours	3	2	3	2	10	20
<1 week	5	1	2	4	10	22
1 week - 1 month	8	5	6	9	4	32
>1 month	-	2	4	2	4	12
Not Available	-	-	1	1	2	4
Total	16	10	16	18	30	90

Regarding new connections, businesses in 50% of very small economies have to wait between 1 week and 1 month for a new water connection. This was also true of electricity and telephones. On the other hand, the other population categories seem to have more extreme cases (more than 1 month or even not available), so there may even be a revealed advantage to smallness. In fact, however, the formal tests suggest no such effect.

Table 4.79 Chi-Square Statistics for New Water Connections

Type of Grouping	Pearson's Chi Square		
	Value	df	Asymp. Sig. (2 sided)
Pop (1, 2-5)	4.507	4	0.342
Pop (1-2, 3-5)	2.244	4	0.689
GDP (1, 2-5)	1.733	4	0.785
GDP (1-2, 3-5)	8.823	4	0.066

Conclusion

The conclusions that we can take from this section are very tentative, given the caveats presented above. There is evidence of a lack of reliability and high cost of water in the small states studied, but, overall, not sufficient to reject the null hypothesis of no cost disadvantage definitively. There is some evidence that connection fees are lower for small economies.

IV. The Costs of Smallness
4. Utilities
4.4 - Fuel

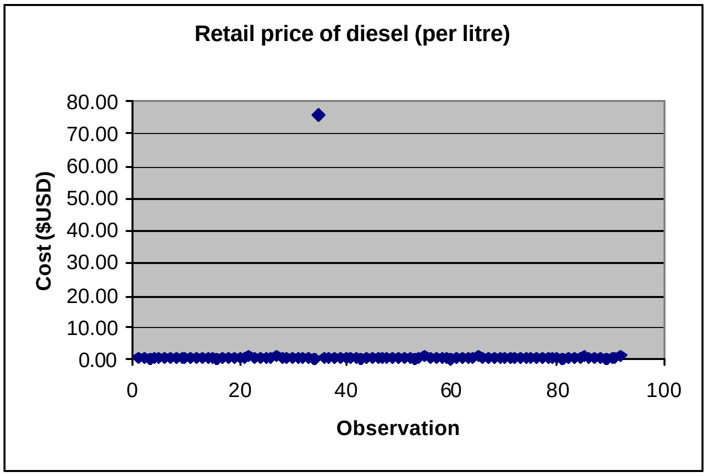
Data Sources

Data on GDP and Population were obtained from the same sources as previously. Information on fuel costs (petrol and diesel) was taken from the Survey. We have valid observations for all countries.

Descriptives

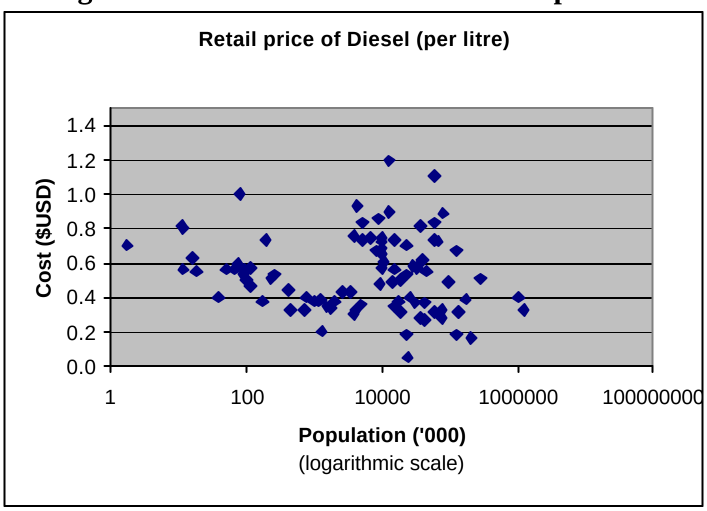
The graph below clearly shows an outlier (Ireland). It seems likely that the value input in the survey (both for diesel and petrol) was in cents instead of dollars. Therefore, we decided to divide both values ($87 and $76) by 100.

Figure 4.30 Retail Price of Diesel



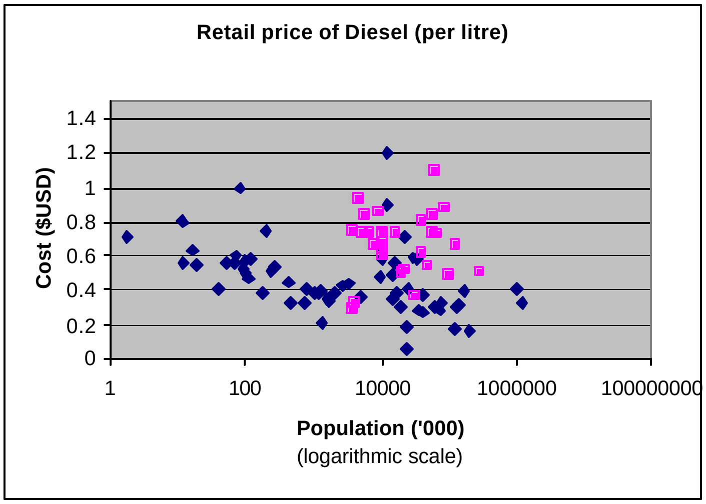
The plot of fuel costs against population seems to suggest that there might be a break point somewhere near 10 million inhabitants.

Figure 4.31 Retail Price of Diesel vs. Population



However, as we have seen in previous graphs, this pattern may also arise because of an OECD effect. Below we identify the OECD countries plus the 3 high-income Asian economies.

Figure 4.32 Retail Price of Diesel (with Sub-Sample highlighted)



Based on the above graph we decided to carry the analysis of size both with the high-income countries and without. There are at least some good reasons for believing that the high-income countries could behave differently in fuel pricing. Distribution costs, taxation and environmental and safety rules all respond to income levels and levels of development, and so could cause qualitative differences in observed fuel prices.

Analysis of Size

For both Diesel and Petrol, we have used the following specification:

$$Ln(Fuel\ Costs) = \alpha_0 + \alpha_1 * Ln(GDPc) + \alpha_2 * Ln(Pop) + \alpha_3 * [Ln(Pop)]^2$$

Alternatively, we excluded the squared term.

Results

The results with the full sample are as follows (population squared was not significant):

Table 4.80 Results for Fuel Costs (full sample)

	Diesel	**Petrol**[1]
constant	**-0.590**	**-0.598**
	-4.545	*-4.635*
LnGDPc	**0.099**	**0.116**
	3.637	*4.273*
LnPop	**-0.026**	**-0.003**
	-1.884	*-0.186*
R-squared	0.17	0.17
obs.	91	91

[1] Fails Functional Form
Venezuela (-) was excluded from both regressions due to normality problems

Without the OECD and the three high-income Asian countries we have 63 observations, which yield

Table 4.81 Results for Fuel Costs
(Sub-sample excluding high income Asia and OECD)

	Diesel[1]	Petrol[2]
constant	**-0.126**	**-0.314**
	-0.892	*-2.170*
LnGDPc	**-0.110**	**-0.025**
	-2.422	*-0.527*
LnPop	**-0.084**	**-0.041**
	-5.269	*-2.488*
R-squared	0.32	0.11
obs.	62	61

[1] Venezuela (-) was excluded due to normality problems.
[2] Ecuador (+) and Venezuela (-) were excluded due to normality problems.

The results from the table 4.81 show that population has a significant negative effect for both fuel prices. GDP per capita is significant only for diesel. It is also interesting to note that excluding Hong Kong, Taiwan, Singapore and the OECD countries reverses the sign of GDP per capita and makes it insignificant for petrol, as well as increasing the size and significance of the population coefficient.

Cost of Smallness

The results below (table 4.82) indicate that the small countries face high penalties on cost of fuel. Such high prices for an important input may well put significant constraint on external competitiveness.

Table 4.82 Cost Deviation from the Median Country (%) for Fuel

Size	Pop ('000)	Diesel	Petrol
Micro	12.13	75.8	31.7
Very Small	197.00	39.1	17.5
Small	4,018.00	8.0	3.8

Conclusion

Although we have excluded the OECD and 3 Asian economies, we believe that these results are solid in identifying the effects of the country smallness on the prices of fuel. Small developing countries apparently have higher retail fuel prices than larger ones.

IV. The Costs of Smallness

5. Passenger Air Travel Costs

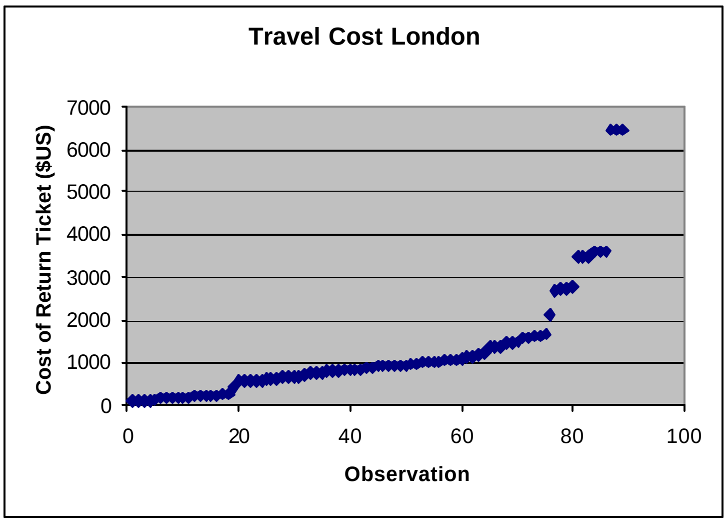
We also collected some data on passenger travel costs. This variable did not feature in the Business Cost Survey, but was obtained from the Commonwealth Secretariat's travel agents. It represents the cost of an economy return ticket from each capital to the respective destination. The remaining variables (Population, Air distances and GDP) are the same as those used for airfreight.

Table 4.83 Missing Observations for Air Travel Costs

London	Tokyo	New York
Anguilla	Anguilla	Anguilla
Swaziland	Japan	Swaziland
UK	Swaziland	UK
	UK	US

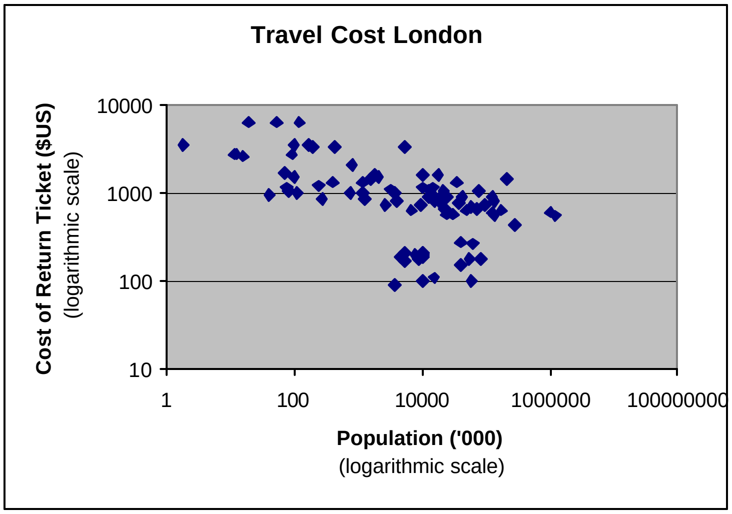
We plotted travel costs for the different destinations. We can observe the extremely high costs corresponding to the Pacific Islands. Most of these flights have to call in Fiji before continuing the journey. Thus, the price is equal to the price of Fiji to the three different destinations plus the flight from Fiji to the other island. (Palau, Micronesia and Marshall Islands were excluded from the regression for London as outliers.)

Figure 4.33 Travel Cost London



Below we have a plot of the travel costs against population. It is clear that there is a decreasing trend of costs with population.

Figure 4.34 Travel Cost London vs. Population



$$Ln(Travel\ Cost) = \alpha_0 + \alpha_1 * Ln(Air\ Distance) + \alpha_2 * Ln(GDPc) + \alpha_3 * Ln(Pop) + \alpha_4 * [Ln(Pop)]^2$$

<u>Results</u>

Table 4.84 Results for Air Travel Costs

	London[1]	**Tokyo**	**New York**[2]
constant	**2.255**	**4.486**	**0.407**
	6.743	*4.895*	*0.515*
LnDistance	**0.641**	**0.455**	**0.871**
	19.741	*4.643*	*10.106*
Ln GDPC	**-0.160**	**-0.285**	**-0.213**
	-7.823	*-7.878*	*-7.001*
LnPop	**-0.116**	**-0.106**	**-0.121**
	-11.340	*-5.643*	*-8.132*
R-squared	0.93	0.58	0.76
obs.	86	88	87

[1] Micronesia (+), Palau (+) and Marshall Islands (+) were deleted due to normality problems. Fails Functional Form.
[2] Canada (+) was deleted due to normality problems. Fails Functional Form.

The squared population term was never significant. We tried to solve the Functional Form problems by including a squared term of GDP per capita. Notwithstanding that this procedure cured the functional form problem for New York, the respective coefficient was insignificant. Therefore, we calculate the cost disadvantages with the output from table 4.84.

Figure 4.35 Travel Cost Tokyo vs. Population

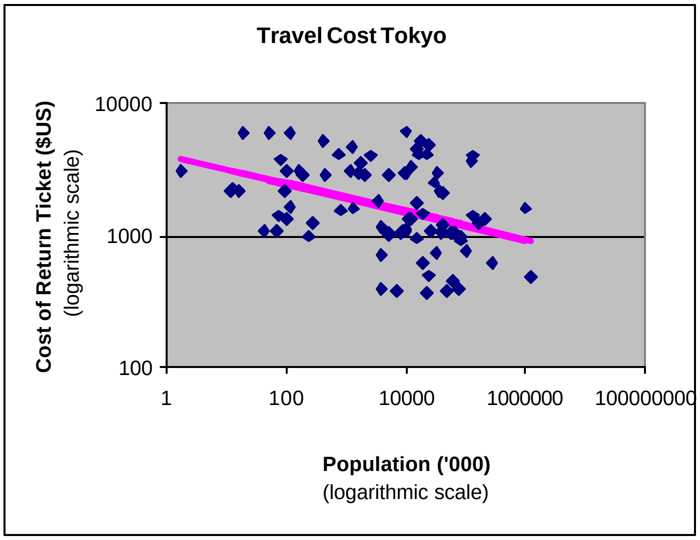


Table 4.85 Cost Deviation from the Median Country (%) for Air Travel Costs

Size	Pop ('000)	London	Tokyo	NY
Micro	12.13	118.0	103.8	125.4
Very Small	197.00	57.7	51.7	60.9
Small	4,018.00	11.2	10.2	11.7

We have also tried to run regressions on duration times for air-trips, but we had no information for the Pacific Islands. The data we have suggest that duration are considerably higher for small economies and adding the Pacific Islands, with their stopovers in Fiji, can only reinforce this view.

High personal airfares and durations are significant for two reasons. For tourism, they are a straight tax on a high proportion of the cost of the business. For manufactures they are not particularly high share of costs, but, particularly the long duration, they are a discouragement to fuller integration with the world economy. If it takes four days travelling to meet face-to-face with a Pacific Island-base collaborator, such a collaboration is pretty costly for a modern executive.

IV. The Costs of Smallness
6. Land

Data Sources

Data on GDP and Population was obtained from the same sources as previously. Information on the cost of land was taken from the Survey.

Data concerns

The main problem with this analysis, as we will see below, is the huge dispersion of land values. Mozambique, especially, deserves mention. It was stated in the survey that the average annual cost of industrial (factory) space was US$2.35 per hectare! This means $0.000235 per square metre! We also had doubts about the values for Mauritius. We infer that the answer 'US$ Rs.1.7 sq.meter/month' is actually expressed in Mauritius Rupees (MUR) and so we converted it to US dollars using the appropriate exchange rate. (The underlining in the response suggests to us that the answer provided included 'Rs.' explicitly, and that the US$ was left in inadvertently.)

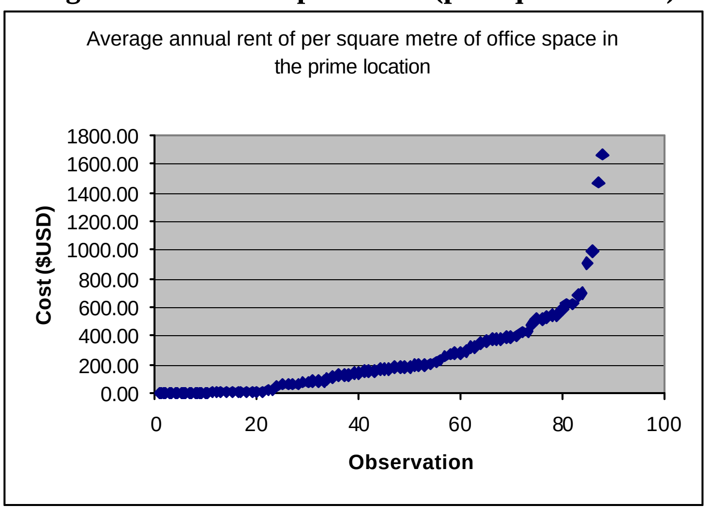
Table 4.86 Missing Observations for Land Costs

	Office	Factory
Missing Obs.	Anguilla Antigua & Barbuda Nauru Tuvalu	Anguilla Nauru Venezuela

Descriptives

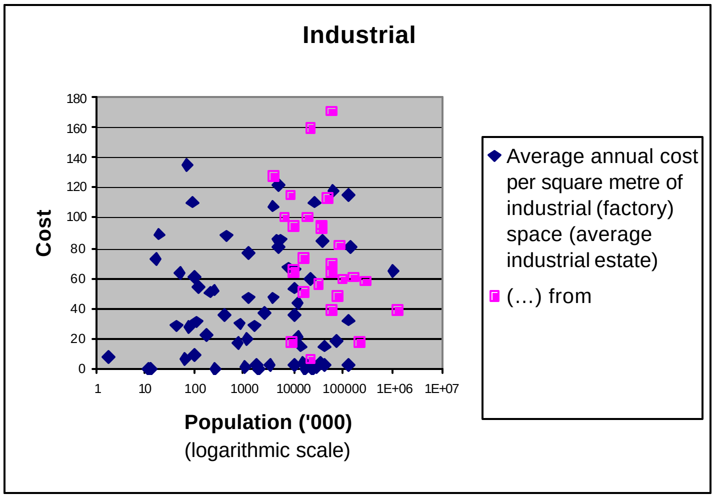
We have information for average annual costs per square metre of industrial (factory) space and average annual rent per square metre of office space.

Figure 4.36 Office Space Rent (per square metre)



However, there is a difference between the EIU questionnaire and the others. The EIU asks for an average value (though stating that a range can be specified), but it presents two boxes to input values (one labelled 'from' and the other 'to'). However, not all countries in the EIU survey have given a range. For around 25 cases there was only a value for 'from', while there was a blank for 'to'. For the countries that stated a range in this questionnaire, we computed a simple average. For the others, the reported value might be interpreted as either the average or a minimum value without maximum. The plot below separates the averages from in which we have only 'from' data.

Figure 4.37 Industrial Space Rent (per square metre) vs. Population (with highlight)



On the whole, we think that the latter represent also the averages, and treat them accordingly. Then, we plotted all the average cost data against population, and observed the correlation between factory and office costs.

Figure 4.38 Industrial Space Rent (per square metre) vs. Population

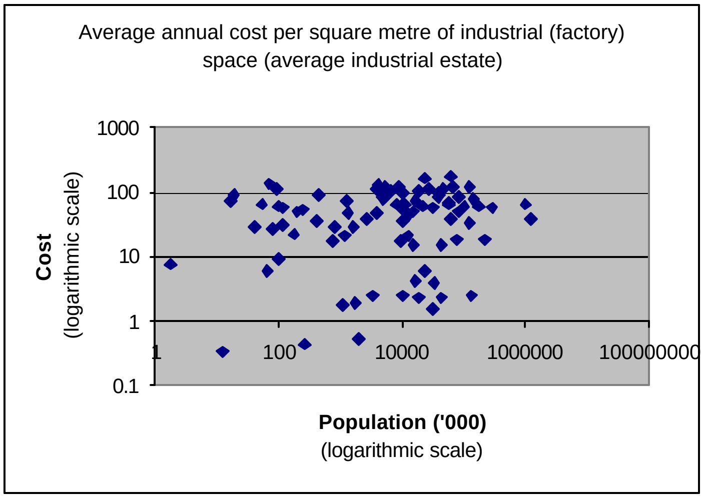
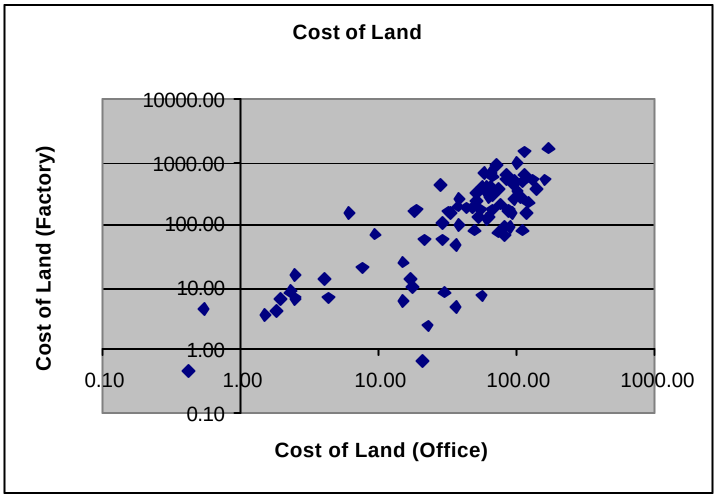


Figure 4.39 Industrial Space Rent vs. Office Space Rent



Analysis of Size

We have used the following specification:

$$\text{Ln(Land Costs)} = \alpha_0 + \alpha_1 \text{*Ln(GDPc)} + \alpha_2 \text{*Ln(Pop)} + \alpha_3 \text{*[Ln(Pop)]}^2$$

Alternatively, we also present the results without the squared term.

Results:

Table 4.87 Results for Land Costs (Office)

	Office
constant	**4.371**
	3.811
LnGDPc	**0.729**
	9.783
LnPop	**-0.399**
	-1.395
LnPop2	**0.035**
	2.068
Wald-test (pop)	16.348[.000]
R-squared	0.51
obs.	86

	Office
constant	**2.532**
	5.075
LnGDPc	**0.710**
	8.946
LnPop	**0.152**
	2.938
-	
-	
R-squared	0.47
obs.	86

Barbados (-) and Mauritius (-) were excluded. The t-statistics use the White's adjusted standard errors to overcome heteroscedasticity.

The turning point of the quadratic form is very low (around the 2nd observation), and so, like the simple log-linear case, it too suggests that smaller economies have lower land costs.

Table 4.88 Results for Land Costs (Factory)

	Factory

	Factory

constant	**2.847**		**constant**	**3.275**
	3.945			*9.585*
LnGDPc	**0.247**		**LnGDPc**	**0.256**
	4.120			*4.520*
LnPop	**0.172**		**LnPop**	**0.040**
	0.948			*1.137*
LnPop2	**-0.009**		**-**	
	-0.806			
Wald-test (pop)	1.553[.460]		-	
R-squared	0.30		R-squared	0.29
obs.	73		obs.	73

Mozambique (-), Barbados (-), Mauritius (-), Tuvalu (-), Lesotho (-), Uruguay (-), Namibia (-), Swaziland (-), Colombia (-), Kenya (-), Sri Lanka (-), Antigua & Barbuda (-), Bangladesh (-), Malaysia (-), Malawi (-) and Cote d'Ivoire (-) were deleted due to normality problems. The t-statistics use the White's adjusted standard errors to overcome heteroscedasticity.

There is no sign of a significant size effect on factory land costs.

Sensitivity Tests

All the countries excluded from the regressions reported in table 4.88 have very low costs of land and most are small. Thus, we have to conclude that land costs increase with size – there is an advantage to small country locations.

Table 4.89 Results for Land Costs (Factory) (failing the normality assumption)

	Factory			Factory
constant	**2.687**		**constant**	**1.862**
	2.238			*2.988*
LnGDPc	**0.596**		**LnGDPc**	**0.586**
	4.751			*4.703*
LnPop	**-0.170**		**LnPop**	**0.083**
	-0.529			*1.235*
LnPop2	**0.016**		**-**	
	0.804			
R-squared	0.22		R-squared	0.21
obs.	89		obs.	89

Both regressions fail Normality and Heteroscedasticity.

Cost of Smallness

Table 4.90 reports the costs of smallness for both the quadratic and log-linear equations for office space in table 4.87. They both suggest advantages but of a slightly different pattern. Our preference is for the quadratic form.

Table 4.90 Cost Deviation from the Median Country (%) for Land Costs

Size	Pop ('000)	Office I	Office II
Micro	12.13	-7.0	-64.0
Very Small	197.00	-.34.7	-45.0

Small	4,018.00	-17.8	-13.0
T point		298.87	-
Closest Country		Barbados	-

Conclusion

Land costs seem to increase with size, offsetting at least partially the cost disadvantages in other areas. Given the known economic attractions of agglomeration, this is perhaps not surprising. Moreover, the Ricardian theory of rent, which attributes to land the surpluses from operations that cannot proceed without it, would suggest that disadvantaged economies would generate very low rents. Thus the positive coefficient on population in this section actually tends to confirm the disadvantages of size noted above. We note the benefits of small size in the summary below, but when we calculate the competitive disadvantages of small size in specific industries below we omit the rent effects because they are endogenous.

IV The Cost of Smallness
7. Conclusion

Tables 4.91 and 4.92 summarise the cost disadvantages faced by three exemplar economies, which we term micro, very small and small. They include some effects based on doubtfully significant co-efficients where we think there is some case for an effect, and zeros where we think there is no case at all. For the most part, however, the recorder effects are significant and, for micro and very small economies, large. In nearly all cases business in small economies faces excess costs. The exceptions are some connection fees for utilities and, for quite different reasons, land. In addition the categorical data suggest that small economies face shortages of skilled labour more frequently than larger ones and have less secure utilities services.

Table 4.91 Summary of Cost Disadvantages

Area of Cost	Micro	Very Small	Small	Comment
Airfreight				
to London	60.3	8.2	-3.1	
from London	*62.1*	*18.9*	*1.3*	Not significant
to Tokyo	85.3	15.2	-2.2	
from Tokyo	*7.1*	*-0.4*	*-1.2*	Not significant
to New York	45.2	1.0	-4.9	
from New York	*37.2*	*3.2*	*-3.1*	Not significant
Seafreight				
to Rotterdam	195.3	67.0	9.3	
from Rotterdam	287.4	100.1	14.6	
to Yokohama	301.5	87.2	10.4	
from Yokohama	251.6	85.0	11.9	
to New York	148.3	44.4	4.5	Fails Functional Form
from New York	145.6	46.7	5.5	
Nominal Wages in dollars				

Construction Worker	65.5	34.3	7.1	
Checkout Operator	43.7	23.6	5.1	
Kitchen Porter	71.1	36.9	7.6	
Bank Clerk/Teller (local)	39.9	21.7	4.7	
Bank Clerk/Teller (foreign)	*30.8*	*17.0*	*3.7*	Not significant
Garage Mechanic	49.6	26.6	5.6	
Payroll Clerk	*8.4*	*4.8*	*1.1*	Not significant
Qualified Teacher	37.1	20.3	4.4	
Bank Manager (local)	53.7	28.6	6.0	
Bank Manager (foreign)	*36.2*	*19.8*	*4.3*	Not significant
General Registered Nurse	61.1	32.2	6.7	
Manufacturing Costs	-	-	-	Not significant
Telephone				
Local Calls	0	0	0	5 zeros excluded
London	97.1	48.7	9.7	Without Pop^2
Tokyo	119.4	58.4	11.3	Without Pop^2
New York	177.6	81.7	14.9	Without Pop^2 and Fails FF
Installation Fee	-32.3	-20.4	-5.2	1 zero excluded
Line Rental	*19.1*	*10.8*	*2.4*	Not significant (2 zeros excl.)
Electricity				
Cost	93.1	47.0	9.4	1 zero excluded
Connection	*-49.9*	*-33.3*	*-9.0*	Not significant (9 zeros excl.)
Water				
Cost	0	0	0	4 zeros excluded
Connection	-77.0	-57.7	-18.1	11 zeros excluded
Fuel				
Diesel	75.8	39.1	8.0	
Petrol	31.7	17.5	3.8	
Air Travel Costs				
London	118.0	57.7	11.2	Fails Functional Form
Tokyo	103.8	51.7	10.2	
New York	125.4	60.9	11.7	Fails Functional Form
Land				
Office	-7.0	-34.4	-17.8	
Factory	0	0	0	

Table 4.92 Summary of the Analysis of the Discrete Data

Area of Cost	Cross Tabulations	Ordered Logits	Comment
Availability of Workers			
Unskilled	Not significant	Significant at 10%	Weak evidence
Semi-skilled	Significant	Significant	Strong evidence
Skilled	Significant	Significant	Strong evidence
Telephone			

New Connection	Significant for GDP	Not Significant	No evidence
Disruption	Significant at 10%	Significant	Good evidence
Repair Times	Not Significant	Not Sginificant	No evidence
Electricity			
Connection	Not Significant	Not Significant	No evidence
Disruption	Significant	Significant	Strong evidence
Water			
Connection	Not Significant	Not Significant	No evidence
Disruption	Significant at 10%	Significant	Good evidence