Comparative advantage, the rank-size rule, and Zipf's law

Hinloopen, J.; van Marrewijk, C.

Publication date
2009

Comparative advantage, the rank-size rule, and Zipf's law[1]

Jeroen Hinloopen[2]

Universiteit van Amsterdam and *Katholieke Universiteit Leuven*

Charles van Marrewijk[3]

University Utrecht

February 2009

Abstract
Using a comprehensive international trade data set we investigate empirical regularities known as Zipf's Law or the rank-size rule for the distribution of the interaction between countries as measured by revealed comparative advantage. Using the recently developed estimator by Gabaix and Ibragimov (2007) we find strong evidence in favor of the rank-size rule along the time, country, and sector dimension for three different levels of data aggregation. The estimated Pareto exponents that characterize the distribution of revealed comparative advantage are stable over time but differ across countries and sectors. These differences are related empirically to country and sector characteristics, including population size, GDP, and factor intensities.

Keywords: Revealed comparative advantage, Balassa index, rank-size rule, Zipf's Law, power law, Pareto distribution

JEL codes: C13, C33, F11, O50, R12

---

[1] Thanks are due to Harry Bowen, Xavier Gabaix, Harry Garretsen, Sidney Resnick, Casper de Vries, and to seminar participants at the Technical University Eindhoven, the Utrecht School of Economics, the Erasmus School of Economics (Rotterdam), the Institute for Advanced Studies (IHS, Vienna), Kingston University (London), and the annual conference of the Royal Economic Society, 2008, Warwick (Coventry). The usual disclaimer applies.
[2] Corresponding author. Correspondence: University of Amsterdam, FEB/ASE, Roetersstraat 11, 1018 WB Amsterdam, The Netherlands; J.Hinloopen@uva.nl; www.fee.uva.nl/io/jhinloopen.
[3] Correspondence: Utrecht School of Economics, Utrecht University, Janskerkhof 12, 3512 BL Utrecht, The Netherlands; J.G.M.vanMarrewijk@uu.nl; www.charlesvanmarrewijk.nl.

"No phenomenon is a phenomenon until it is an observed phenomenon"

Niels Bohr

1. Introduction

It is often the revelation of empirical regularities that leads to the development of new theories. For instance, Frank Auerbach's observation that city sizes are distributed regularly has led to the development of a body of literature that explains this observed pattern (Nitsch, 2005), the observation by Jan Tinbergen that gravity laws empirically rule international trade patterns initiated a literature that grasps the economics behind the "gravity equation" (Frankel, 1998), and Robert Engle's observation that in many time series the variance changes over time triggered several theories that explain this observed conditional heteroskedasticity (Semmler, 1994).

In this paper a new empirical regularity is documented: the obedience of revealed comparative advantage to the rank-size rule. The notion of revealed comparative advantage was first introduced by Liesner (1958) and operationalized by Balassa (1965) with his concomitant index. The latter is defined as a country's exports in some sector as a fraction of national exports, divided by world exports in that sector as a fraction of world exports. Whenever the Balassa index exceeds unity, a comparative advantage is 'revealed' for the particular country in the particular sector. On average about one third of all sectors display such a revealed comparative advantage, although this percentage varies considerably across countries (Hinloopen and van Marrewijk, 2001).

The rank-size rule holds if a log-linear relationship exists between the value of some phenomenon and its rank in the related sample. In the special case of a slope equal to one the rank-size rule is labeled "Zipf's law", named after the Harvard linguistic professor George Kingsley Zipf (Zipf, 1949). In this paper we carefully document the fact that the rank-size rule applies whenever a comparative advantage is revealed by the Balassa index, and that in a substantial, but minority, of cases Zipf's law holds. Our analysis is based on a comprehensive data set that is obtained by merging two sets obtained from the Center for International Data at the University of California, Davis (see Feenstra *et al*., 1997, and Feenstra, 2000). It consists in particular of observations on bilateral trade flows for 747 4-digit sectors, 166 countries, covering the years 1970 through 1997, yielding a total of slightly less than 18.4 million positive observations. This allows for a thorough and systematic empirical analysis of the observed phenomenon along three different dimensions: over time, across countries, and across sectors.

Although applications of the rank-size rule abound, this paper deviates from the existing empirical literature in three main respects. First, countries rather than cities are our unit of observation. In this respect we follow Rose (2005) who argues that the size distribution of countries is similar to that for cities.[4] Second, the distribution of the interaction between economic centers rather than the size distribution as such is our focus of analysis. The innovation here is to use the empirical notion of revealed comparative advantage for capturing the interaction between countries rather than the commonly used gravity equation. Third, we use the estimation procedure recently introduced by Gabaix and Ibragimov (2007) for estimating the Pareto exponent in the rank-size rule equation, as this procedure eliminates the bias inherent to traditional estimators.

We find strong evidence in favor of the rank-size rule for international trade flows that reflect a comparative advantage. Although we focus on reporting the estimates of the Pareto exponents of the Balassa index distribution, some further empirical explorations suggest that these exponents are systematically related to country-specific characteristics and that they differ systematically across sectors in case these are grouped according to factor-intensities. Borrowing the words of Rose (2005, p. 11), these findings add to what amounts to be "an intriguing puzzle for future theoretical work."

The remainder of the paper is organized as follows. Section 2 briefly reviews the rank-size rule and its application to the Balassa index. Section 3 discusses some estimation issues. Section 4 contains our empirical findings regarding the applicability of the rank-size rule (and, in particular, Zipf's law) to revealed comparative advantage. Section 5 discusses these findings and presents some further empirical explanations. Section 6 concludes.

2. The rank-size rule and comparative advantage

The rank-size rule states that the frequency of occurrence of some event P as a function of the concomitant rank i that follows from this frequency is a power-law function: $P_i \sim 1/i^{\phi}$. This means that the distribution of the event is approximately Pareto. In case ϕ equals one the rank-size rule is referred to as Zipf's law. Phenomena in economics and finance abound which exhibit the implied heavy-tailedness in the data (i.e. $\phi > 0$), see e.g. Mandelbrot (1963), Janssen and de Vries (1991), Gabaix (1999), Gabaix *et al*. (2003), Axtell (2001), and Nitsch (2005).

It appears that revealed comparative advantage also complies to the rank-size rule. The concept of revealed comparative advantage is widely used empirically to identify a country's weak and strong export sectors (for recent applications see e.g. Porter (1990), Amiti (1999), Fertö and Hubbard (2003), or Svaleryd and Vlachos, 2005). In particular, let $X_{i,t}^{j}$ be the value of exports from country $i \in I$ for sector $j \in J$ in period $t \in T$. Then $X_{i,t} = \sum_j X_{i,t}^{j}$ is the

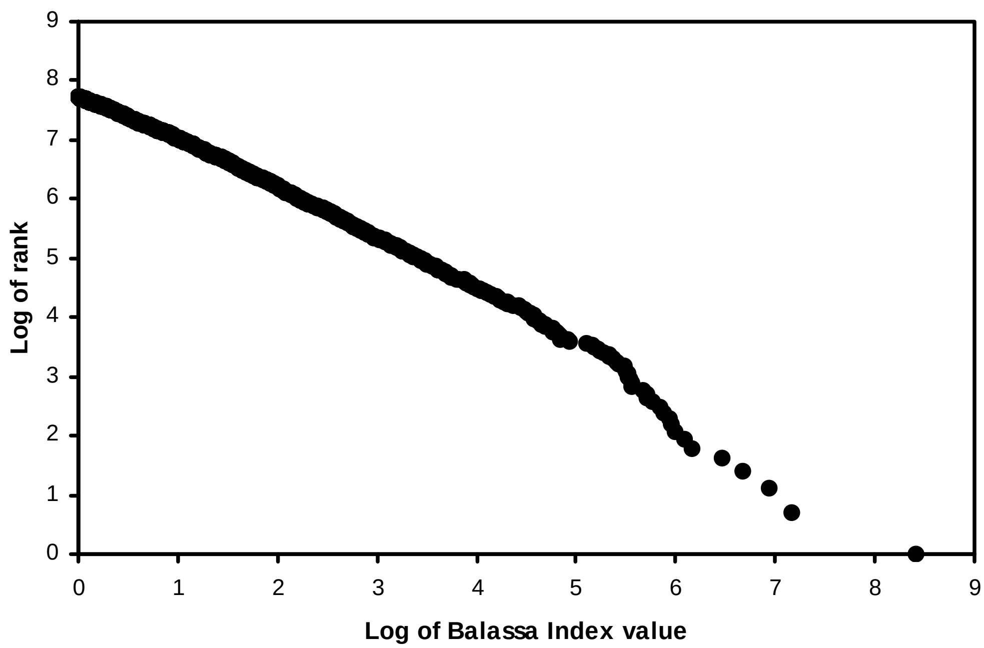
[4] See also Okuyama, Takayasu, and Takayasu (1999) and Axtell (2001) for similar evidence regarding firm size distributions.

value of exports from country i in period t, $X_t^j = \sum_i X_{i,t}^j$ is the total value of exports for sector j in period t, $X_t = \sum_i \sum_j X_{i,t}^j$ is the total value of exports for all countries and sectors in that period, and the Balassa index for country i and sector j in period t is defined as:

$$(1) \qquad B_{i,t}^j = \frac{X_{i,t}^j / X_{i,t}}{X_t^j / X_t}, \quad i \in I,\ j \in J, t \in T.$$

If $B_{i,t}^j > 1$, country i is said to have a revealed comparative advantage in the production of commodity j in time period t as its export share for product j is larger than the concomitant export share in the group of reference countries I.[5]

Figure 1. Empirical relation between the Balassa index of revealed comparative advantage and the related rank. The Figure is based on 1% randomly drawn Balassa index values from the entire sample of 221,430 observations.



In Figure 1 the log of the Balassa index is plotted against the log of the concomitant rank (the figure is constructed using a random sample of 1,000 observations; appendix A contains the description of the data used throughout the paper), whereby the largest Balassa index value is given rank 1. This picture serves as a first illustration that the rank-size rule applies to the

[5] Hillman (1980) identifies a sufficient condition for the Balassa index to measure comparative advantage proper in that an increase in exports yields an increase in the Balassa index. Throughout the paper the analysis is restricted to those observations that meet this "Hillman condition", which amounts to disgarding 0.25% of all observations (see Hinloopen and Van Marrewijk (2008) for further details).

upper tail of the empirical distribution of Balassa index values. Indeed, observations with a Balassa index above unity are grouped almost perfectly on a straight line. It is precisely these observations that reveal a comparative advantage.

Whether or not Zipf's law holds in particular has been of interest in the related literature on city size distributions and, recently, the size distribution of firms (Luttmer, 2005, Melitz, 2003). Meanwhile the applicability of the rank-size rule to city-size distributions is undisputed but to date there is no agreement as to the particular validity of Zipf's law. Some argue that the Pareto exponent does not assume unity quite regularly (Rosen and Resnick (1980), Brakman, Garretsen, and van Marrewijk (2001), and Soo, 2005), others stress that in most studies the hypothesis that the estimated Pareto exponent equals one cannot be rejected (Krugman (1996), Gabaix (1999), and Gabaix and Ioannides, 2004). Given the scope of our dataset we are quite confident that our findings regarding the applicability of the rank-size rule and, in particular, Zipf's law, to revealed comparative advantage, are representative for the underlying mechanisms.

3. Estimating a power-law coefficient

The two most commonly used methods for estimating the Pareto exponent ϕ are OLS (often referred to as the *Zipf regression*) and employing the *Hill estimator* (Hill (1975), see also Gabaix and Ioannides, 2004). For large sample sizes the estimated Pareto exponent in the Zipf regression tends with probability one to the true value of ϕ. For small samples, however, the estimate is biased and inefficient. Moreover, the reported standard errors grossly underestimate the true standard errors. Alternatively, under the null hypothesis of a perfect power law, the Hill estimator is the maximum likelihood estimator of ϕ. But the properties of the Hill estimator in finite samples are also worrisome as the bias can be very high in small samples and the associated computed standard errors considerably underestimate the true standard errors as a result of this bias. As the rate of convergence can be arbitrary slow (Embrechts *et al*., 1997) the estimator also requires very large samples sizes. Moreover, the choice as to the number of order statistics to be included is problematic in view of the bias – variance tradeoff (see Beirlant *et al*. (2004) and the citations therein).

Various estimators have been developed to address these issues, but these have not led to a consensus solution of the problem (see e.g. Embrechts *et al*., 1997, Beirlant *et al*., 1999, and Feuerverger and Hall, 1999). Recently however Gabaix and Ibragimov (2007) provide an elegant and effective solution for the estimation of Pareto exponents: an unbiased estimate is obtained when using OLS in:

(2) $$\ln\left(rank - \frac{1}{2}\right) = a - b\ln(size) .$$

Accordingly, all that is needed is to shift the rank by ½. Gabaix and Ibragimov (2007) show further that the standard error of the so-estimated Pareto exponent is asymptotically equal to $\sqrt{(2/n)}b$. Also, using OLS in (2) is more robust to deviations from the exact power law formulation (in the sense of Hall, 1982) than is the Hill estimator (see also Ibragimov and Phillips (2004), and Phillips, 2007).

However, given that routine data typically carry up to 10 percent contamination (Hampel *et al*., 1986) OLS estimates in (2) are likely to be biased. This calls for including as many observations as possible as Gabaix and Ibragimov (2007) show that this bias is inversely related with the sample size. Accordingly, we employ as low as possible a cut-off value for identifying sectors with a revealed comparative advantage. That is, include all observations with a Balassa index of at least one. By doing so we also split the sample at a theoretically meaningful point: include all observations that relate to a revealed comparative advantage.

In sum, we report only the Gabaix-Ibragimov estimates and their associated standard errors for all observations that reveal a comparative advantage.[6]

4. Empirical results

Since the Balassa index carries three dimensions (time, country, and sector), we can empirically investigate the size of the Pareto exponent for these three dimensions. *Zipf's law* is said to apply in the particular dimension if the estimated slope coefficient of equation (2) does not differ significantly from unity; in all other cases the *rank-size rule* is said to apply given that the estimated coefficient is statistically significant. Observe that there is a sub-dimension for the Balassa index with respect to the degree of data aggregation. After eliminating erroneously classified observations, our data set effectively distinguishes 66 2-digit sectors, 225 3-digit sectors, and 419 4-digit sectors (see Appendix A).[7] All three identified dimensions are thus considered at three different levels of data aggregation.

4.1 Dimension I: time

To examine the applicability of the rank-size rule along the time dimension of revealed comparative advantage the following equation is estimated for all Balassa index values above unity:

[6] In all cases we also calculated the Zipf regression and the Hill estimator. These are available from the authors upon request.

[7] To be included an observation (i) has a related export value of at least 5,000 US \$, (ii) meets the Hillman condition, and (iii) is at least one. Further, regressions with 10 or fewer observations were dismissed. The smallest number of observations for any regression is 47.

$$(3) \qquad \ln\left[rank\left(BI_{i,t}^{j}\right)-\frac{1}{2}\right] = \alpha_{0,t} + \alpha_{1,t}\ln\left[BI_{i,t}^{j}\right], \quad t = 1970,\ldots,1997.$$

Table 1 contains the summary statistics for the several per-year estimates (details of the estimates for all individual years are listed in Table B1 in Appendix B), and the pooled regression whereby all years are taken together. The estimated Pareto exponents vary little per year and the goodness-of-fit is very high. At the 3-digit level, for example, the estimates range from 0.896 to 0.981 and the goodness-of-fit is between 98.3 and 99.4 per cent.

Table 1. Pooled Gabaix-Ibragimov regression of the slope parameter in (3) and summary statistics based on the 28 concomitant per-year estimates.

	2-digit	3-digit	4-digit
average	0.958	0.930	0.870
median	0.949	0.922	0.860
stand. dev.	0.043	0.025	0.020
min	0.882	0.896	0.849
max	1.031	0.981	0.915
average # observations	1,911	5,080	7,908
pooled	0.955	0.928	0.869
$\overline{R}^2$	0.968	0.991	0.991
# observations	53,503	142,231	221,430

The distribution of the slope estimates is illustrated in Figure 2. Note that all estimated slope coefficients are highly statistically significant and that in all years the goodness-of-fit indicator is very high, especially for estimates at the 2- and 3-digit level of data aggregation.[8] We thus conclude:[9]

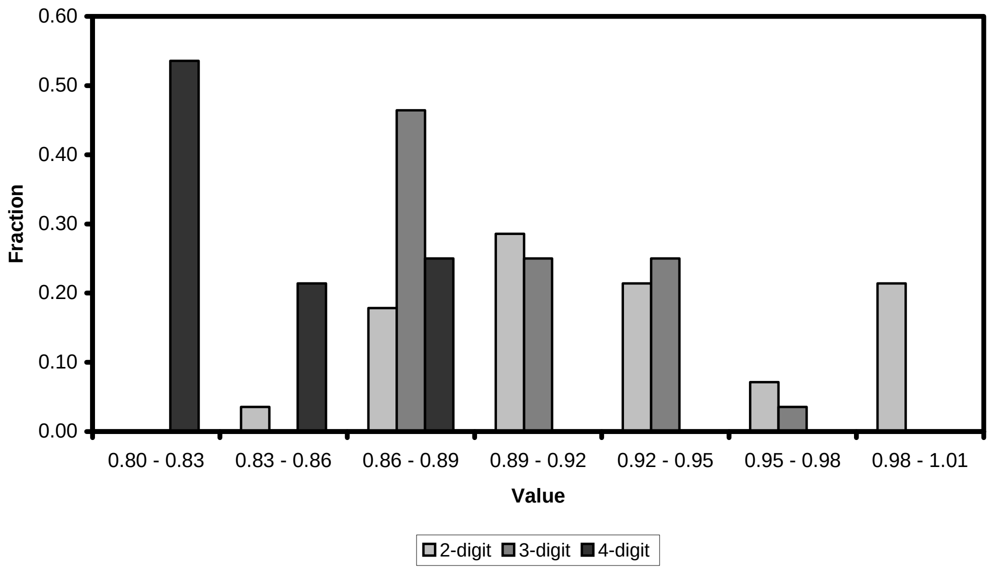
Empirical result 1

The distribution of revealed comparative advantage for the years 1970 through 1997 as measured with the Balassa index at either the 2-digit, 3-digit, or 4-digit level of data aggregation follows the rank-size rule, whereby for the respective levels of data aggregation the mean estimated power coefficient is 0.96 (0.04), 0.93 (0.03), and 0.87 (0.02).

[8] An obvious additional test for examing the validity of the linearity of the relationship is to include a quadratic term in (2), as in Black and Henderson (2003). However, as illustrated by Gabaix and Ioannides (2004), it is very likely that the estimated coefficient of this quadratic term will turn out to be statistically significant in situations where the underlying data are obtained from a data generating process that is known to comply perfectly to Zipf's law.

[9] The number in brackets are the standard deviation of the several estimates.

Figure 2. Distribution of the estimated per-year slope coefficients at the 2-digit, 3-digit, and 4-digit level of data aggregation.



To examine whether Zipf's law applies in particular a *t*-test is conducted, which leads to:

Empirical result 2

The distribution of revealed comparative advantage for the years 1970 through 1997 as measured with the Balassa index at either the 2-digit, 3-digit, or 4-digit level of data aggregation follows Zipf's law in, respectively, 69%, 24% and 0% of all sample years.

4.2 Dimension II: country

The country dimension is captured by the following equation:

$$\ln\left[rank\left(BI_{i,t}^{j}\right) - \frac{1}{2} \right] = \beta_{0,i} + \beta_{1,i} \ln\left[BI_{i,t}^{j}\right], \quad i = 1, \Lambda\ ,166. \tag{4}$$

Table 2 provides summary statistics of the related estimates for the different levels of aggregation whereas details of the individual estimates for the 166 countries are in Table A2 in Appendix B.

Table 2. Summary statistics of the 166 per-country Gabaix-Ibragimov estimates of the slope parameter in (4).

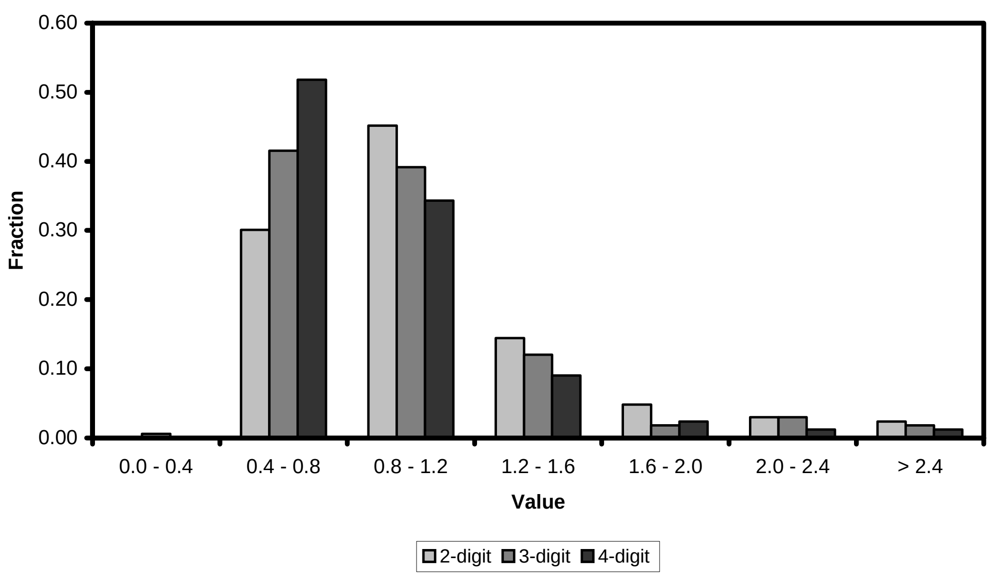
	2-digit	3-digit	4-digit
average	1.048	0.964	0.890
median	0.939	0.839	0.790
stand. dev.	0.474	0.432	0.386
min	0.498	0.366	0.444
max	3.710	3.326	2.948
average # observations	322	857	1,334

All estimated slope coefficients are statistically significant, and in almost all cases the goodness-of-fit is very high (the median value of the latter is 89.5, 92.7, and 94.2 per cent at the 2-digit, 3-digit, and 4-digit level, respectively). Hence:

Empirical result 3

The distribution of revealed comparative advantage for the 166 sample countries as measured with the Balassa index at either the 2-digit, 3-digit, or 4-digit level of data aggregation follows the rank-size rule, whereby for the respective levels of data aggregation the mean estimated power coefficient is 1.05 (0.47), 0.96 (0.43), and 0.89 (0.39).

Figure 3. Distribution of the estimated per-country slope coefficients at the 2-digit, 3-digit, and 4-digit level of data aggregation.



As is clear from Table 2 and illustrated in Figure 3, the estimated Pareto exponents vary considerably between countries. At the 3-digit level, for example, the maximum estimate of 3.326 for Germany is almost ten times as high as the minimum estimate of 0.366 for Comoros. Although the goodness-of-fit also varies more than before, in all cases the explanatory power remains high.[10] This suggests that an individual country exhibits its own characteristic distribution of revealed comparative advantage which might be related to country factors (an issue taken up further in Section 5 below). Finally, considering the *t*-tests for examining whether or not Zipf's law applies leads us to observe:

Empirical result 4

The distribution of revealed comparative advantage for the 166 sample countries as measured with the Balassa index at either the 2-digit, 3-digit, or 4-digit level of data aggregation follows Zipf's law in, respectively, 38%, 20% and 13% of all cases.

4.3 Dimension III: sectors

The sector dimension is examined with the following equation:

$$\text{(5)} \qquad \ln\left[rank\left(BI_{i,t}^{j}\right)-\frac{1}{2}\right] = \gamma_0^j + \gamma_0^j \ln\left[BI_{i,t}^{j}\right], \quad j = 1,\Lambda\ ,J,$$

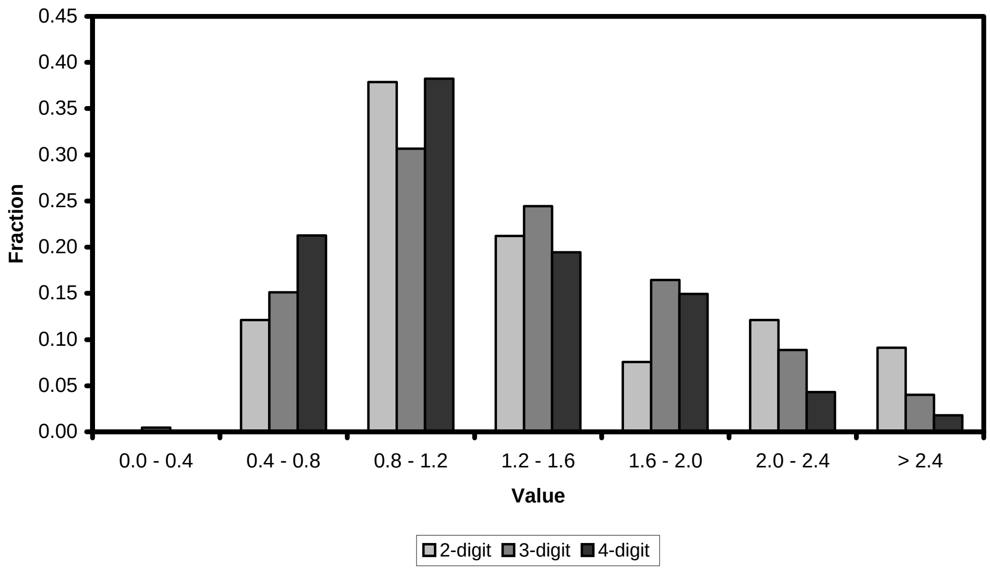
whereby J refers to the number of sectors for the particular level of data aggregation. At the 2-digit level there are 66 different sectors, at the 3-digit level 225 sectors are distinguished, and 419 sectors are identified at the 4-digit level. Table 3 provides the summary statistics for the sector estimates and Table B3 in Appendix B contains the details of all sector estimates separately. Figure 4 illustrates the distribution of the estimated sector exponents.

Table 3. Summary statistics of the 66 2-digit, the 225 3-digit sectors, and the 419 4-digit per-sector Gabaix-Ibragimov estimates of the slope parameter in (5).

	2-digit	3-digit	4-digit
average	1.411	1.347	1.198
median	1.225	1.253	1.101
stand. dev.	0.618	0.542	0.469
min	0.646	0.394	0.415
max	3.113	3.420	3.282
average # observations	811	632	528

[10] This holds *a fortiori* if the number of observations is taken into account; on average a relatively low goodness-of-fit is related to a limited number of observations. For instance, the regression for Germany is based on 2,942 observations while the estimate for Comoros uses 142 observation only (Table B2).

Figure 4. Distribution of the estimated per-sector slope coefficients at the 2-digit, 3-digit, and 4-digit level of data aggregation.



We arrive at similar conclusions as for the country dimension. The estimated Pareto exponents vary considerably between sectors even for the same level of data aggregation. At the 3-digit level, for instance, the minimum estimate is 0.394 for sector 264 ("jute and other textile bast fibres, nes, raw/processed") and the maximum estimate is 3.420 for sector 784 ("parts and accessories of 722, 781, 782, 783 [relating to motor vehicles]"). Further, the goodness-of-fit for the individual estimates remains high: the median is 94.3, 94.9, and 94.7 per cent at the 2-digit, 3-digit, and 4-digit level, respectively. We thus conclude:

Empirical result 5

The distribution of revealed comparative advantage for the 66 2-digit sectors, the 225 3-digit sectors, and the 419 4-digit sectors as measured with the Balassa index follows the rank-size rule, whereby for the respective levels of data aggregation the mean estimated power coefficient is 1.40 (0.62), 1.34 (0.54) and 1.20 (0.47).

The considerable variation across sectors in the estimated power-law coefficients is also examined further in Section 5 below. For now we observe that for a substantial number of sectors Zipf's law applies:

Empirical result 6

The distribution of revealed comparative advantage for the 66 2-digit sectors, the 225 3-digit sectors, and the 419 4-digit sectors as measured with the Balassa index follows Zipf's law in, respectively, 16%, 17% and 26% of all cases.

5. Discussion

We thus far have seen that revealed comparative advantage always complies to the rank-size rule, and that a modest share of the estimated Pareto exponents does not differ significantly from unity. In addition, the estimated Pareto exponents do not differ from year to year, but they do vary between countries and between sectors.

5.1 Characteristics of estimated country exponents

The significant differences in estimated Pareto exponents between countries suggests that these estimates are related to country characteristics. To give further weight to this conjecture we relate the estimated Pareto exponents to country characteristics that can be expected to affect the value of the Balassa index, including country size (measured by GDP and number of inhabitants), export size, openness (exports as a fraction of GDP), and export breadth (measured as the number of sectors in which a comparative advantage is obtained). At the same time, the results in Section 4.1 imply that the estimated power coefficient are stable over time. This leads us to consider the following equation:

$$\gamma_{it} = \beta_i + \beta_t + \beta_1 \gamma_{it-1} + \beta_2 \ln(GDP_{it}) + \beta_3 \ln(POP_{it}) + \beta_4 \ln(EXP_{it}) + \beta_5 \frac{EXP_{it}}{GDP_{it}} + \beta_6 \ln\left(EXPS_{it}\right) + \varepsilon_{it}, \tag{6}$$

whereby γ_{it} is the estimated value of the power-law exponent for country i in year t, β_i is a country-specific fixed effect, β_t is a time-specific fixed effect, *POP* is the population size (measured in 1000 persons), *EXP* is the value of exports (measure in 1000 US \$), *EXPS* is the number of export sectors, and ε is assumed to be *iid* with zero mean and constant variance.[11] However, in (6) the estimated power-law exponent $\gamma_{it\text{-}1}$ depends on β_i which leads to biased estimates. To solve this problem first differences are considered:

$$\begin{aligned} \gamma_{it} - \gamma_{it-1} &= \alpha_t + \beta_1 (\gamma_{it-1} - \gamma_{it-2}) + \beta_2 \left(\ln(GDP_{it}) - \ln(GDP_{it-1})\right) \\ &\quad + \beta_3 \left(\ln(POP_{it}) - \ln(POP_{it-1})\right) + \beta_4 \left(\ln(EXP_{it}) - \ln(EXP_{it-1})\right) \\ &\quad + \beta_5 \left(\frac{EXP_{it}}{GDP_{it}} - \frac{EXP_{it-1}}{GDP_{it-1}}\right) + \beta_6 \left(\ln\left(EXPS_{it}\right) - \ln(EXPS_{it-1})\right) + \varepsilon_{it} - \varepsilon_{it-1}, \end{aligned} \tag{7}$$

[11] See Appendix A for the several data sources.

where $\alpha_t = \beta_t - \beta_{t-1}$. This creates yet another problem as $\gamma_{it\text{-}1}$ is correlated with $\varepsilon_{it\text{-}1}$. Following Anderson and Hsiao (1982) this is restored by using $\gamma_{it\text{-}2}$ as an instrument for ($\gamma_{it\text{-}1}$-$\gamma_{it\text{-}2}$).

Table 4. Panel estimates for explaining cross country differences in estimated Pareto exponents (6); dependent variable: first differences estimated Pareto exponents.*

	2-digit	3-digit	4-digit
time fixed effect	-0.00	0.04	0.04
	(0.07)	(3.40)	(4.02)
lagged power-law coefficient	0.03	-0.03	-0.04
	(1.48)	(4.10)	(4.59)
country size (GDP)	0.66	0.29	0.21
	(2.59)	(4.34)	(3.51)
country size (POP)	0.10	-0.60	-0.30
	(0.09)	(2.22)	(1.20)
export size	-0.54	-0.04	-0.08
	(11.09)	(3.53)	(8.15)
export breadth	-0.56	0.08	0.05
	(8.22)	(4.49)	(3.35)
openness	345.38	65.42	80.61
	(3.57)	(2.56)	(3.04)
# observations	3058	3113	3100
$\overline{R}^2$	0.05	0.02	0.03

* *t*-values are within brackets.

The results of the panel estimates are in Table 4, where the estimated power-law coefficients are from the regressions of Section 4.2. [12] Considering these results prompts two observations. First, although the specification in (7) is not obtained from a formally derived first-order condition, the strong significance of most estimated coefficients does suggest that the cross-country variation of the estimated Pareto exponent is related to country-specific characteristics. Second, at the 3-digit and 4-digit level the estimated panel coefficients are quite comparable. At the 2-digit level the export breadth variable is still significant but with

[12] In total the panel study is based on 13,944 annually estimated Pareto exponents at the country level (namely 28 years × 166 countries × 3 levels of aggregation). Appendix B contains all these point estimates of γ.

an opposite sign. This is due to the related drop in the number of distinguished sectors which alienates the export breadth variable in the 2-digit case from those in which either 3-digit or 4-digit export flows are considered. We conclude:

Empirical result 7

The estimated power coefficient of the rank-size rule for revealed comparative advantage as measured with the Balassa index differs significantly across countries. These differences are related to the following country-specific characteristics: country size, export size, export breadth, and openness.

5.2 Characteristics of estimated sector exponents

Differences across sectors regarding the estimated Pareto exponent can be structured effectively according to the core analysis of international trade: factor intensity. To that end we use the factor intensity classification of the International Trade Center, the joint UNCTAD/WTO organization. In particular we have the following five broad factor intensity categories (within brackets the number of 3-digit sectors belonging to the particular category; note that there are 5 sectors that are not classified; see Hinloopen and van Marrewijk, 2008):[13]

A. *Primary products* (83); including meat, dairy, cereals, fruit, coffee, sand, minerals, oil, natural gas, iron ore, and copper ore.
B. *Natural-resource intensive products* (21); including leather, cork, wood, lime, precious stones, pig iron, copper, aluminum, and lead.
C. *Unskilled-labor intensive products* (26); including various textiles and garments, clothing, glass, pottery, ships, furniture, footwear, and office supplies.
D. *Human-capital intensive products* (43); including synthetic colors, pigments, perfumes, cosmetics, rubber and tires, tubes, pipes, various types of steel and iron, cutlery, televisions, radios, cars, watches, and jewellery.
E. *Technology intensive products* (62); including various chemicals, medicaments, plastics, engines, generators, machines, tools, pumps, telecommunications and photo equipment, optical equipment, and aircraft.

Table 5 contains the average estimated power-law coefficients for these factor intensity categories. It appears that the estimated Pareto exponents are about equal for primary products and natural-resource intensive products. Similarly, they are about equal for technology-intensive products and human-capital intensive products. Moreover, the estimated

[13] We first made some adjustments to get from the SITC Rev. 3 codes to the codes in Feenstra *et al*. (1997); see http://people.few.eur.nl/vanmarrewijk/eta for further details.

exponents tend to be lower for primary products than for unskilled-labor intensive products, which in turn tend to be lower than that for human-capital intensive products. Indeed, according to a two-sided *t*-test of mean differences the following ordering applies:

$$\gamma_{\text{primary}} = \gamma_{\text{natural-resource}} < \gamma_{\text{unskilled-labor}} < \gamma_{\text{human-capital}} = \gamma_{\text{technology}}$$

Table 5. Pair wise *t*-tests for equality of the mean of estimated power-law coefficients across sectors (3-digit).

	A	B	C	D	Mean estimate	*t*-value for H_0: $\overline{\gamma} = 1$
A	*				0.94	2.05
B	0.86	*			1.01	0.12
C	5.16	3.25	*		1.37	4.73
D	8.59	5.93	2.73	*	1.67	13.71
E	13.10	8.15	4.51	1.37	1.81	8.42

In addition, we have tested whether Zipf's law applies for the five different factor intensity categories. Although the mean value of the estimated power-law coefficient for natural-resource intensive products does not differ from that for primary products, it is only for the former product category that Zipf's law applies in particular (see Table 5). We thus conclude:

Empirical result 8

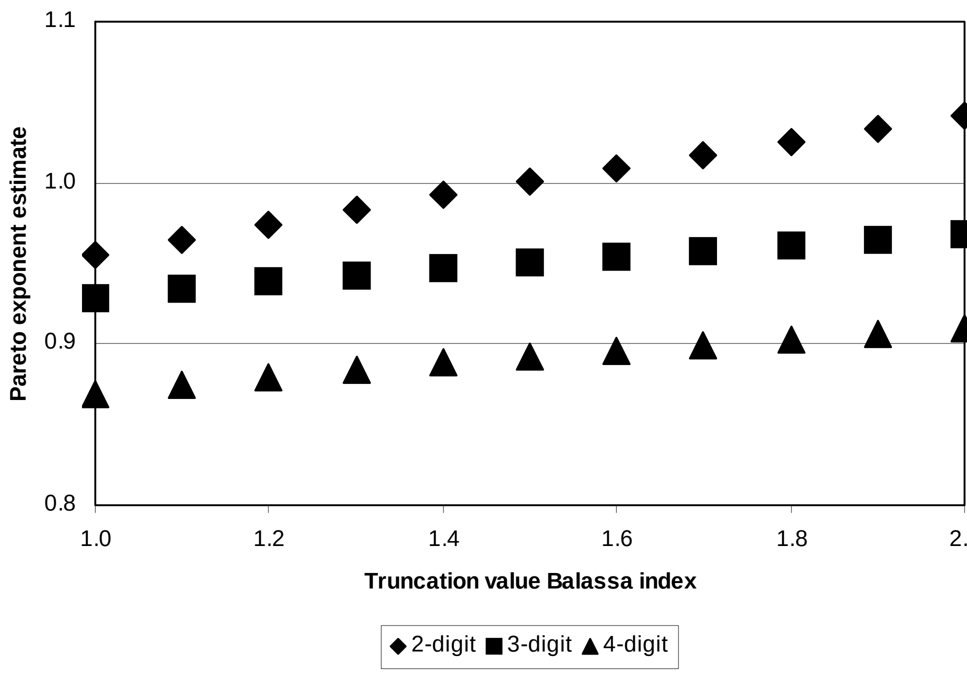
The estimated power coefficient of the rank-size rule for revealed comparative advantage as measured with the Balassa index differs significantly across sectors ordered according to factor intensities. For all sectors the rank-size rule applies; for natural-resource intensive sectors Zipf's law holds in particular.

6. Conclusions

The analysis of the evolution of a country's apparent strong export sectors as indicated by a revealed comparative advantage using the Balassa index (or its monotonic transformations) is complicated by the lack of a characterization of the distribution of this index. Using a comprehensive international trade data set we reveal an empirical regularity of revealed comparative advantage that alleviates some of this complication: its obedience to the rank-size rule along the time, country, and sector dimension. Our results are obtained using the estimator recently developed by Gabaix and Ibragimov (2007) which yields unbiased estimates absent data contamination. As our analysis is based in total on 1,295 estimated

Pareto exponents we are quite confident that our findings are representative for the phenomenon documented here.

Figure 5. Estimated power-law coefficients for different truncation values at the 2-digit, 3-digit, and 4-digit level of data aggregation.



Eeckhout (2004) shows the estimated Pareto exponent for city size distributions to be very sensitive to the sample composition. Using a higher truncation point for cities to be included into the sample leads to higher power law estimates. As there is no natural truncation point for city sizes Eeckhout's observation questions the applicability of the rank-size rule to city size distributions. For the Balassa index on the other hand the subsample of entries above one naturally presents itself. Thus, in Figure 5 the (absolute value of the) estimated Pareto exponent is plotted for higher truncation values.[14] The observation of Eeckhout (2004) also applies here although the estimated power coefficient is much less sensitive to the truncation value compared to the city size distribution. This supports the robustness of our findings.

The evidence presented in this paper calls for theoretical explorations, possibly along the lines of the recent explanations for city size distributions (Gabaix (1999), Brakman *et al.*

[14] Note that the fraction of industries included drops with higher truncation values. For the 2-digit, 3-digit, and 4-digit level of data aggregation going from a truncation value of one to two comes with a drop in this fraction respectively from 24.5% to 14.4%, from 24.4% to 13.9% and from 27.2% to 16.3%. Increasing the truncation value even further yields estimates that are based on a disproportionately small part of the sample.

(2001), and Eeckhout, 2004) or firms (Luttmer, 2005). The variation in estimated Pareto exponents across countries and sectors suggests that these theories should take into account country and/or sector characteristics. This is confirmed by the panel estimates that document the empirical relation between estimated Pareto exponents and country specific characteristics, including country size, openness, and export breadth. The market size variables point to theories of imperfect competition in which the home-market effect, the competition for large markets, and location effects through intra-industry trade linkages play a prominent role.[15]

At the sectoral level the estimated Pareto exponents differ systematically across sectors classified along factor intensity. This finding points to Heckscher-Ohlin-Samuelson type international trade theories which are rooted in factor abundance. To the extent that countries differ in technology, physical- and human capital abundance, or (un)skilled labor abundance, each of which affects the sectoral composition of revealed comparative advantage, this should also be taken into consideration when analyzing the estimated Pareto exponent at the country level. We leave that and the theoretical explorations for future research.

References

Amiti, M. (1999), "Specialization patterns in Europe," *Weltwirtschaftliches Archiv*, 135: 573 – 593.

Anderson, T.W. and C. Hsiao (1982), "Formulation and estimation of dynamic models using panel data", *Journal of Econometrics*, 18: 47 – 82.

Axtell, R.L. (2001), "Zipf distribution of U.S. firm sizes", *Science* 293: 1818 – 1820.

Balassa, B. (1965), "Trade liberalization and 'revealed' comparative advantage", *The Manchester School of Economic and Social Studies* 33: 92 – 123.

Beirlant, J., G. Dierckx, Y. Goegebeur, and G. Matthys (1999), "Tail index estimation and an exponential regression model", *Extremes* 2: 177 – 200.

Beirlant, J., Y. Goegebeur, J. Teugels, and J. Segers (2004), *Statistics of extremes*, Wiley Series in Probability and Statistics, John Wiley & Sons Ltd., Chichester.

Black, D. and V. Henderson, 2003, "Urban evolution in the USA", *Journal of Economic Geography*, 3: 343 – 372.

Brakman, S., H. Garretsen, and C. van Marrewijk (2001), *An introduction to geographical economics: trade, location, and growth*, Cambridge University Press, Cambridge, U.K.

Eeckhout, J. (2004), "Gibrat's law for all cities," *American Economic Review* 94: 1429-1451.

[15] See also the related recent theories explaining the empirical success of the gravity equation in Evenett and Keller (2002) and the geographical economics theories on the endogenous determination of location in Fujita, Krugman, and Venables (1999).

Embrechts, P., C. Kluppelberg, and T. Mikosch (1997), *Modelling extremal events for insurance and finance*, Springer, New York.

Evenett, S.J., and W. Keller (2002), “On theories explaining the success of the gravity equation”, *Journal of Political Economy* 110: 281-316.

Feenstra, R., R. Lipsey, and H. Bowen (1997), “World trade flows, 1970 – 1992, with production and tariff data”, NBER Working Paper No. 5910.

Feenstra, R. C. (2000), *World trade flows; 1980 – 1997*, Institute of Governmental Affairs, University of California, Davis.

Fertö, I., and L.J. Hubbard (2003), “Revealed comparative advantage and competitiveness in Hungarian Agri-food sectors”, *World Economy*, 26: 247 – 259.

Feuerverger, A., and P. Hall (1999), “Estimating a tail exponent by modelling departure from a Pareto distribution”, *Annals of Statistics* 27: 760-781.

Frankel, J. (ed.), 1998, *The regionalization of the world economy*, NBER.

Fujita, M., P.R. Krugman, and A.J. Venables (1999), *The spatial economy: cities, regions, and international trade*, MIT Press, Cambridge, MA.

Gabaix, X. (1999), “Zipf’s Law for cities”, *Quarterly Journal of Economics* 114: 739 – 767.

Gabaix, X., P. Gopikrishnan, V. Plerou, and H.E. Stanley (2003), “A theory of power-law distributions in financial market fluctuations”, *Nature* 423: 267 – 270.

Gabaix, X., and Y. Ioannides (2004), “The evolution of city size distributions”, in: J.V. Henderson and J.-F. Thisse (eds.), *Handbook of Regional and Urban Economics IV: Cities and Geography*, North-Holland: 2341 – 2378.

Gabaix, X., and R. Ibragimov (2007), “Rank-1/2: a simple way to improve the OLS estimation of tail exponents”, NBER Technical Working Paper No. 342.

Hall, P. (1982), On some simple estimates of an exponent of regular variation”, *Journal of the Royal Statistical Society, Series B* 44: 37-42.

Hampel, F.R., E.M. Ronchetti, P.J. Rousseeuw, and W.A. Stahel (1986), *Robust statistics: the approach based on influence functions*, New York: Wiley.

Hill, B.M. (1975), “A simple approach to inference about the tail of a distribution”, *Annals of Statistics* 3: 1163 – 1174.

Hillman, A.L. (1980), “Observations on the relation between ‘revealed comparative advantage’ and comparative advantage as indicated by pre-trade relative prices”, *Weltwirtschaftliches Archiv* 116: 315 – 321.

Hinloopen, J., and C. van Marrewijk (2001), “On the empirical distribution of the Balassa index”, *Weltwirtschaftliches Archiv* 137: 1 – 35.

Hinloopen, J., and C. van Marrewijk (2008), “The empirical relevance of the Hillman condition for revealed comparative advantage”, *Applied Economics* 40: 2313 – 2328.

Ibragimov, R., and P.C.B. Phillips (2004), "Regression asymptotics using martingale convergence methods", Cowles Foundation, Discussion Paper No. 1473.

Janssen, D., and C.G. de Vries (1991), "On the frequency of large stock returns: putting booms and busts into perspective", *Review of Economics and Statistics* 73: 18 – 32.

Krugman, P.R. (1996), "Confronting the mystery of urban hierarchy", *Journal of the Japanese and International Economies* 10: 399 – 418.

Liesner, H.H. (1958), "The European Common Market and British industry", *Economic Journal* 68: 302 – 316.

Luttmer, E.G.J. (2005), "Growth and the size distribution of firms", *Quarterly Journal of Economics*, forthcoming.

Maddison, A. (2003), *The world economy: historical statistics*, OECD, Paris, France.

Mandelbrot, B. (1963), "The variation of certain speculative prices", *Journal of Business* 36: 394 – 419.

Melitz, M (2003), "The impact of trade on intra-industry reallocations and aggregate industry productivity", *Econometrica*, 71(6): 1695 – 1725.

Nitsch, V. (2005), "Zipf zipped", *Journal of Urban Economics* 57: 86 – 100.

Okuyama, K., M. Takayasu, and H. Takayasu (1999), "Zipf's Law in income distribution of companies", *Physica A* 269: 125 – 131.

Phillips, P.C.B. (2007), "Regression with slowly varying regressors", *Econometric Theory*, forthcoming.

Porter, M. (1990), *The competitive advantage of nations*, MacMillan, London.

Rose, A.K. (2005), "Cities and countries", NBER Working Paper No. 11762.

Rosen, K.T., and M. Resnick (1980), "The size distribution of cities", *Journal of Urban Economics* 8: 165 – 186.

Semmler, W. (1994), Editor, *Business cycles: theory and empirical evidence*, Kluwer Academic Publishers.

Soo, K.T. (2005), "Zipf's law for cities: a cross-country investigation", *Regional Science and Urban Economics* 35: 239 – 263.

Svaleryd, H. and J. Vlachos (2005), "Financial markets, the pattern of industrial specialization and comparative advantage: evidence from OECD countries", *European Economic Review*, 49: 113 – 144.

World Bank (2005), *World Development Indicators*, Washington D.C., U.S.A.

Zipf, G.K. (1949), *Human behavior and the principle of least effort*, Addison-Wesley, Cambridge, MA.

Appendix A Data

Two separate data sets compiled by the Center for International Data, University of California, Davis (CID/UCD), were merged, the first covering the years 1970 through 1993 (see Feenstra, Lipsey and Bowen, 1997) and the second covering the years 1980 through 1997 (see Feenstra, 2000). For the overlapping years, the data from the latter source are used. The data set contains bilateral trade flows between 183 trading partners, including n.e.s. (not elsewhere specified) regions for trade flows that could not be classified further than within a broad geographical region (such as "Middle East", or "North Africa"), an "Areas n.e.s." region for trade flows that cannot be attributed to any country or to any of the used broad geographical regions but that do come from a well-defined geographical region, and an "Unknown Partner" category for trade flows that could not be attributed at all due to various reasons (see Feenstra, 2000). This leaves a sample of 165 genuine countries

The bilateral trade flows are decomposed into 1,249 sectors, comprising 747 genuine 4-digit sectors, based on SITC (Standard International Trade Classification), revision 2. The remaining 502 sectors refer to aggregates at the 1-, 2-, or 3-digit level, and a "Non-identified products" category. The 4-digit subset contains 60.39 % of all trade, the 3-digit subset covers 99.46 % of all trade, and the 2-digit subset comprises 99.67 % of all trade.

The data were first compiled by Statistics Canada and made available through the CID/UCD (see Feenstra, 2000). The former makes use of various sources (according to Statistics Canada 87% of all trade flows is based on independent sources of *both* imports and exports, while 98% is based on reports of at least one side of trade), yielding a rather complete coverage of world trade flows. The CID/UCD transforms the data such that trade flows for all years, all countries, and all industry groups are consistent and presented in a unified manner. Each observation in the raw data consists of four entries: importing country, exporting country, sector, and size of the trade flow (in 1,000 US $). The data are thus classified according to the importing country. This is *not* to say that the data are based on import sources only, as explained above. After merging the two separate datasets a second dataset is created by "inverting" the data, in that all trade is classified according to the exporting country.

The income and population data are taken from the World Bank Development Indicators 2005 except for Taiwan. Those data are based on own calculations using the Maddison (2003) data set and complementary material taken from the website of the National Statistics Office of Taiwan (http://eng.stat.gov.tw/mp.asp?mp=5).

Table A1 Sample countries

Country Nr.	Name	Country Nr.	Name
1	South Africa	91	Belize
2	Algeria	92	Falkland Isl
3	Liby Arab Jm	93	French Guiana
4	Morocco	94	Guyana
5	Western Sahara	95	Panama
6	Sudan	96	Surinam
7	Tunisia	98	Israel
8	Egypt	99	Japan
10	Cameroon	100	Bahrain
11	Central Afr. Rep.	101	Cyprus
12	Chad	102	Iran
13	Congo	103	Iraq
14	Gabon	104	Jordan
16	Angola	105	Kuwait
17	Br.Ind.Oc.Tr	106	Lebanon
18	Burundi	107	Oman
19	Comoros	108	Qatar
20	Zaire	109	Saudi Arabia
21	Benin	110	Fm Dem Yemen
22	Eq. Guinea	111	Syrn Arab Rp
23	Ethiopia	112	Untd Arab Em
25	Djibouti	113	Turkey
26	Gambia	114	Fm Yemen
27	Ghana	115	Yemen
28	Guinea	117	Afghanistan
29	Cote D'ivoire	118	Bangladesh
30	Kenya	119	Bhutan
31	Liberia	120	Brunei
32	Madagascar	121	Myanmar (Burma)
33	Malawi	122	Cambodia
34	Mali	123	Sri Lanka
35	Mauritania	124	Hong Kong
36	Mauritius	125	India
37	Mozambique	126	Indonesia (Incl Macau)
38	Niger	127	Korea Rp (South)
39	Nigeria	128	Laos P.Dem.R
40	Guinea-Bissau (Incl Cape Verde)	129	Malaysia
41	Reunion	130	Maldives
42	Rwanda	131	Nepal
43	St.Helena	132	Pakistan
44	Senegal	133	Philippines
45	Seychelles	134	Singapore
46	Sierra Leone	135	Thailand
47	Somalia	136	Taiwan
48	Zimbabwe	137	China
49	Togo	138	Korea D P Rp (North)
50	Uganda	139	Mongolia
51	Untd Rp Tanzania	140	Vietnam
52	Burkina Faso	142	Belgium-Lux.
53	Zambia	143	Denmark (Incl Faroe Islds)

Table A1 Sample countries (continued)

Nr.	Name	Nr.	Name
55	Canada	144	France
56	Usa	145	Germany
57	Bermuda	146	Greece
58	Greenland	147	Ireland
59	St Pierre Miqu	148	Italy
60	Argentina	149	Netherlands
61	Bolivia	150	Portugal
62	Brazil	151	Spain
63	Chile	152	United Kingdom
64	Colombia	154	Austria
65	Ecuador	155	Finland
66	Mexico	156	Iceland
67	Paraguay	157	Norway
68	Peru	158	Sweden
69	Uruguay	159	Switzerland
70	Venezuela	161	Gibraltar
72	Costa Rica	162	Malta
73	El Salvador	164	Albania
74	Guatemala	165	Bulgaria
75	Honduras	166	Czechoslovakia
76	Nicaragua	167	Fm German Dm Rp (East)
78	Bahamas	168	Hungary
79	Barbados	169	Poland
80	Cayman Islds	170	Romania
81	Cuba	172	Fm Yugoslavia
82	Dominican Rp	173	Fm Ussr
83	Guadeloupe (Incl Martinique)	174	Australia
84	Haiti	175	New Zealand
85	Jamaica	176	Solomon Islds
86	Neth Antilles	177	Fiji
87	St Kitts Nev (Incl Dominica, Montserrat, St Luca,St Vinct, Grenada)	178	Kiribati (Incl Solomon Islds, Tonga, Tuvalu)
88	Trinidad-Tobago	179	New Caledonia (Incl Fr Polynesia, Vanuata)
89	Turks Caicos Isl	180	Papua N.Guinea

Other geographical regions in data set (nes = not elsewhere specified)

Nr.	Name	Nr.	Name
9	North Africa nes	141	Asia Cpe nes
15	Ceuca nes	153	Eec nes
24	Fr.So.Ant.Tr	160	Efta nes
54	Other Africa nes	163	Other Eur nes
71	Laia nes	171	Fm Eur Cpe nes
77	Cacm nes	181	Oth. Oceania nes
90	Caribbean nes	182	Areas nes
97	Rest America nes	183	Unknown Partner
116	Middle East nes		

Appendix B Power law estimates

Table B1 GI Pareto exponent estimates of Rank-Size rule; BI > 1 , time dimension

	2-digit trade flows				3-digit trade flows				4-digit trade flows			
year	estimate	st error	fit	# obs	estimate	st error	fit	# obs	estimate	st error	fit	# obs
1970	1.022	0.036	0.961	1,646	0.968	0.021	0.983	4,261	0.907	0.015	0.982	7,081
1971	1.016	0.036	0.961	1,636	0.960	0.021	0.986	4,304	0.894	0.015	0.983	7,231
1972	1.012	0.035	0.957	1,703	0.961	0.020	0.988	4,479	0.898	0.015	0.984	7,351
1973	1.031	0.035	0.960	1,719	0.978	0.021	0.990	4,533	0.915	0.015	0.987	7,591
1974	1.023	0.034	0.964	1,789	0.981	0.020	0.991	4,838	0.904	0.014	0.989	8,134
1975	1.008	0.034	0.966	1,793	0.964	0.020	0.992	4,779	0.899	0.014	0.990	8,017
1976	0.999	0.033	0.963	1,802	0.958	0.019	0.992	4,826	0.883	0.014	0.990	7,889
1977	1.017	0.034	0.960	1,794	0.973	0.020	0.991	4,850	0.891	0.014	0.990	7,947
1978	0.979	0.031	0.966	1,961	0.926	0.018	0.991	5,147	0.852	0.013	0.989	8,338
1979	0.977	0.032	0.970	1,906	0.931	0.018	0.993	5,073	0.854	0.013	0.991	8,272
1980	0.962	0.030	0.970	2,015	0.928	0.018	0.994	5,379	0.855	0.013	0.992	8,212
1981	0.953	0.030	0.968	2,017	0.924	0.018	0.992	5,405	0.850	0.013	0.988	8,129
1982	0.943	0.030	0.975	2,021	0.920	0.018	0.992	5,266	0.857	0.014	0.989	7,985
1983	0.946	0.030	0.966	1,935	0.920	0.018	0.992	5,067	0.861	0.014	0.987	7,648
1984	0.938	0.030	0.972	1,917	0.915	0.018	0.992	5,106	0.858	0.014	0.989	7,681
1985	0.951	0.031	0.964	1,937	0.922	0.018	0.989	5,045	0.859	0.014	0.989	7,653
1986	0.945	0.031	0.966	1,869	0.912	0.018	0.989	5,008	0.856	0.014	0.990	7,527
1987	0.940	0.031	0.964	1,891	0.909	0.018	0.990	5,057	0.860	0.014	0.990	7,575
1988	0.947	0.030	0.966	1,928	0.915	0.018	0.990	5,054	0.850	0.014	0.992	7,547
1989	0.954	0.030	0.967	1,957	0.922	0.018	0.990	5,235	0.849	0.014	0.993	7,852
1990	0.927	0.029	0.969	1,981	0.908	0.018	0.992	5,347	0.850	0.013	0.992	7,946
1991	0.910	0.029	0.965	1,973	0.907	0.018	0.990	5,307	0.861	0.014	0.993	8,007
1992	0.905	0.028	0.966	2,026	0.909	0.017	0.991	5,455	0.859	0.013	0.994	8,274
1993	0.895	0.028	0.975	2,090	0.900	0.017	0.992	5,540	0.851	0.013	0.995	8,353
1994	0.882	0.028	0.971	2,053	0.896	0.017	0.991	5,409	0.857	0.013	0.993	8,230
1995	0.897	0.028	0.966	2,042	0.902	0.017	0.992	5,479	0.869	0.013	0.994	8,344
1996	0.911	0.028	0.970	2,045	0.912	0.017	0.993	5,560	0.871	0.013	0.993	8,419
1997	0.947	0.030	0.966	2,057	0.925	0.018	0.991	5,422	0.883	0.014	0.993	8,197
all	0.955	0.006	0.968	53,503	0.928	0.003	0.991	142,231	0.869	0.003	0.991	221,430

Table B2 GI Pareto exponent estimates of Rank-Size rule; BI > 1, country dimension

	2-digit trade flows				3-digit trade flows				4-digit trade flows			
country	estimate	st error	fit	# obs	estimate	st error	fit	# obs	estimate	st error	fit	# obs
1	1.220	0.079	0.955	477	1.100	0.045	0.926	1,217	0.968	0.029	0.915	2,250
2	0.876	0.138	0.812	81	0.940	0.091	0.873	213	1.007	0.084	0.935	289
3	0.958	0.174	0.884	61	1.081	0.140	0.910	119	1.109	0.134	0.916	138
4	0.904	0.062	0.937	426	0.839	0.035	0.966	1,151	0.860	0.026	0.961	2,106
5	0.567	0.087	0.869	85	0.537	0.065	0.908	137	0.461	0.057	0.870	132
6	0.744	0.067	0.744	246	0.681	0.046	0.870	436	0.509	0.028	0.918	656
7	0.980	0.075	0.889	338	0.958	0.043	0.910	997	0.884	0.031	0.955	1,628
8	1.249	0.099	0.955	320	1.051	0.050	0.948	877	0.928	0.035	0.936	1,421
10	0.943	0.075	0.926	314	0.820	0.048	0.917	582	0.751	0.038	0.920	774
11	0.849	0.086	0.841	193	0.654	0.055	0.896	286	0.596	0.045	0.887	344
12	0.583	0.080	0.836	107	0.574	0.063	0.926	164	0.552	0.057	0.959	189
13	1.138	0.154	0.907	109	0.852	0.093	0.879	166	0.689	0.069	0.920	202
14	1.010	0.125	0.714	130	0.722	0.078	0.870	171	0.544	0.054	0.825	202
16	1.043	0.122	0.861	146	0.868	0.079	0.950	243	0.895	0.070	0.963	330
17	0.713	0.082	0.953	151	0.645	0.060	0.906	232	0.579	0.059	0.919	192
18	0.625	0.074	0.823	142	0.598	0.056	0.879	224	0.589	0.047	0.901	316
19	0.498	0.079	0.650	80	0.366	0.043	0.839	142	0.530	0.057	0.886	171
20	1.020	0.084	0.879	294	0.738	0.048	0.876	483	0.702	0.039	0.881	649
21	0.690	0.063	0.847	237	0.686	0.044	0.940	486	0.588	0.034	0.978	593
22	0.628	0.077	0.732	134	0.567	0.053	0.861	233	0.537	0.047	0.901	261
23	0.717	0.061	0.860	280	0.725	0.045	0.932	510	0.574	0.030	0.889	748
25	0.794	0.082	0.972	188	0.838	0.053	0.975	501	0.695	0.040	0.958	610
26	0.617	0.064	0.768	188	0.581	0.045	0.888	338	0.530	0.037	0.889	411
27	0.825	0.071	0.905	267	0.720	0.051	0.922	402	0.703	0.045	0.852	498
28	0.637	0.073	0.845	154	0.612	0.055	0.896	247	0.531	0.045	0.937	274
29	0.832	0.069	0.915	294	0.778	0.047	0.912	560	0.729	0.035	0.868	868
30	0.910	0.064	0.950	400	0.758	0.036	0.972	890	0.821	0.029	0.953	1,626
31	0.738	0.077	0.755	185	0.560	0.049	0.815	259	0.502	0.045	0.797	251

Table B2	2-digit trade flows				3-digit trade flows				4-digit trade flows			
country	estimate	st error	fit	# obs	estimate	st error	fit	# obs	estimate	st error	fit	# obs
32	0.895	0.065	0.943	383	0.680	0.036	0.966	725	0.765	0.033	0.914	1,081
33	0.606	0.058	0.850	221	0.595	0.039	0.935	462	0.542	0.028	0.924	755
34	0.759	0.060	0.931	316	0.691	0.043	0.963	517	0.718	0.038	0.917	711
35	0.577	0.078	0.631	111	0.548	0.053	0.853	216	0.597	0.047	0.905	316
36	0.561	0.062	0.852	164	0.775	0.049	0.926	497	0.798	0.037	0.958	931
37	1.013	0.070	0.863	423	0.814	0.039	0.885	882	0.766	0.029	0.915	1,391
38	0.585	0.084	0.837	98	0.577	0.059	0.961	194	0.702	0.067	0.901	221
39	1.077	0.150	0.893	103	0.923	0.101	0.858	166	0.775	0.074	0.948	220
40	0.762	0.066	0.846	267	0.644	0.040	0.912	511	0.532	0.030	0.971	626
41	0.529	0.076	0.859	98	0.633	0.060	0.939	225	0.615	0.050	0.964	304
42	0.661	0.091	0.806	105	0.568	0.061	0.815	176	0.507	0.045	0.793	254
43	0.738	0.085	0.926	152	0.678	0.050	0.937	372	0.736	0.057	0.951	338
44	0.702	0.053	0.818	353	0.719	0.037	0.924	756	0.683	0.029	0.895	1,080
45	0.757	0.079	0.912	185	0.602	0.044	0.915	376	0.565	0.039	0.909	426
46	0.868	0.083	0.804	217	0.703	0.051	0.901	385	0.588	0.039	0.927	461
47	0.645	0.062	0.919	214	0.649	0.049	0.927	356	0.444	0.032	0.893	382
48	0.851	0.067	0.935	321	0.806	0.041	0.954	785	0.725	0.031	0.967	1,102
49	0.580	0.059	0.801	194	0.556	0.038	0.949	432	0.671	0.045	0.972	451
50	0.657	0.073	0.886	164	0.650	0.056	0.927	274	0.579	0.041	0.897	398
51	0.848	0.061	0.874	391	0.692	0.035	0.911	794	0.663	0.028	0.875	1,134
52	0.682	0.076	0.817	163	0.618	0.053	0.906	272	0.537	0.040	0.911	358
53	0.593	0.085	0.774	98	0.676	0.057	0.916	281	0.796	0.050	0.955	509
55	1.732	0.099	0.885	608	1.567	0.054	0.905	1,672	1.322	0.035	0.932	2,893
56	2.250	0.109	0.930	848	2.372	0.068	0.931	2,470	2.336	0.052	0.942	3,965
57	0.731	0.082	0.881	160	0.752	0.061	0.918	301	0.648	0.055	0.915	276
58	0.578	0.077	0.804	114	0.530	0.048	0.746	247	0.462	0.037	0.801	308
59	0.696	0.071	0.893	191	0.661	0.046	0.932	418	0.561	0.035	0.921	527
60	0.934	0.063	0.840	445	0.981	0.037	0.928	1,381	0.836	0.024	0.934	2,378
61	0.863	0.069	0.855	312	0.648	0.040	0.899	529	0.622	0.033	0.922	712
62	1.206	0.073	0.873	541	1.139	0.042	0.934	1,485	0.959	0.028	0.938	2,423

Table B2	2-digit trade flows				3-digit trade flows				4-digit trade flows			
country	estimate	st error	fit	# obs	estimate	st error	fit	# obs	estimate	st error	fit	# obs
63	1.007	0.075	0.826	361	0.938	0.047	0.923	811	0.827	0.029	0.944	1,595
64	0.974	0.071	0.947	374	1.162	0.052	0.969	984	1.025	0.037	0.974	1,545
65	0.917	0.094	0.790	191	0.785	0.053	0.832	433	0.862	0.046	0.966	700
66	1.666	0.115	0.930	419	1.358	0.054	0.963	1,279	1.166	0.033	0.969	2,452
67	0.842	0.061	0.870	377	0.734	0.041	0.905	631	0.646	0.030	0.911	907
68	0.866	0.065	0.845	356	0.830	0.040	0.846	841	0.683	0.025	0.944	1,459
69	1.033	0.065	0.893	506	0.881	0.036	0.910	1,170	0.736	0.023	0.907	1,991
70	1.222	0.162	0.897	114	1.173	0.091	0.909	329	1.031	0.060	0.960	601
72	1.036	0.071	0.910	429	1.124	0.047	0.965	1,131	0.925	0.032	0.970	1,670
73	0.985	0.067	0.963	435	1.061	0.046	0.982	1,044	0.899	0.030	0.970	1,757
74	0.988	0.064	0.920	472	1.009	0.041	0.964	1,205	0.919	0.029	0.981	2,057
75	0.952	0.072	0.897	354	0.834	0.043	0.942	755	0.784	0.031	0.950	1,283
76	0.884	0.060	0.869	437	0.878	0.041	0.945	901	0.844	0.033	0.954	1,344
78	1.012	0.109	0.857	173	0.824	0.062	0.914	357	0.716	0.049	0.919	424
79	0.928	0.067	0.935	379	0.947	0.042	0.934	995	0.897	0.035	0.947	1,345
80	0.834	0.089	0.908	176	0.734	0.054	0.904	366	0.616	0.047	0.953	346
81	0.719	0.073	0.906	192	0.792	0.054	0.928	427	0.654	0.035	0.912	686
82	0.882	0.069	0.918	327	0.837	0.045	0.890	702	0.765	0.032	0.913	1,144
83	0.639	0.068	0.722	175	0.737	0.049	0.881	456	0.773	0.040	0.957	734
84	0.976	0.072	0.866	364	0.878	0.041	0.921	908	0.768	0.031	0.889	1,229
85	0.824	0.072	0.898	264	0.866	0.049	0.969	613	0.801	0.034	0.983	1,078
86	0.968	0.151	0.790	82	0.759	0.075	0.892	207	0.647	0.050	0.739	330
87	1.036	0.078	0.933	351	0.887	0.043	0.943	858	0.832	0.034	0.961	1,229
88	1.118	0.117	0.825	184	1.024	0.068	0.915	449	0.813	0.043	0.926	721
89	0.605	0.086	0.853	98	0.586	0.057	0.927	208	0.738	0.073	0.950	206
91	0.820	0.070	0.936	278	0.823	0.045	0.954	678	0.731	0.032	0.957	1,052
92	0.636	0.073	0.852	153	0.621	0.050	0.942	307	0.584	0.044	0.951	356
93	0.735	0.075	0.880	193	0.661	0.046	0.927	405	0.650	0.042	0.943	483
94	0.664	0.058	0.861	263	0.615	0.041	0.853	454	0.605	0.034	0.945	644
95	0.919	0.063	0.899	421	0.908	0.043	0.946	884	0.824	0.032	0.974	1,294

Table B2	2-digit trade flows				3-digit trade flows				4-digit trade flows			
country	estimate	st error	fit	# obs	estimate	st error	fit	# obs	estimate	st error	fit	# obs
96	0.864	0.079	0.914	237	0.656	0.050	0.879	344	0.555	0.037	0.936	451
98	1.377	0.091	0.948	462	1.256	0.049	0.973	1,333	1.115	0.033	0.972	2,339
99	2.557	0.173	0.919	436	2.229	0.073	0.923	1,886	1.973	0.055	0.908	2,594
100	0.924	0.116	0.971	127	0.927	0.072	0.948	334	0.870	0.051	0.969	593
101	1.062	0.071	0.928	449	1.023	0.043	0.937	1,137	0.822	0.026	0.955	1,945
102	1.021	0.138	0.922	109	1.034	0.092	0.898	251	1.004	0.069	0.938	424
103	0.762	0.118	0.911	84	0.792	0.080	0.920	196	0.766	0.058	0.975	348
104	0.837	0.059	0.954	403	0.835	0.036	0.986	1,078	0.879	0.031	0.948	1,631
105	1.100	0.134	0.925	135	1.028	0.076	0.940	368	0.940	0.062	0.960	453
106	1.240	0.073	0.980	575	1.107	0.040	0.956	1,528	0.939	0.027	0.958	2,368
107	0.817	0.141	0.911	67	0.899	0.094	0.955	183	1.153	0.118	0.968	192
108	0.997	0.137	0.840	106	1.011	0.103	0.886	192	0.833	0.097	0.899	148
109	1.025	0.163	0.874	79	1.033	0.103	0.880	200	1.094	0.102	0.943	229
110	1.237	0.245	0.904	51	1.119	0.145	0.889	119	1.033	0.101	0.889	210
111	1.055	0.097	0.953	235	0.984	0.056	0.943	627	0.903	0.042	0.966	914
112	0.965	0.131	0.880	108	1.033	0.083	0.947	308	1.004	0.080	0.982	314
113	1.142	0.078	0.919	431	1.110	0.042	0.910	1,417	0.965	0.026	0.942	2,792
114	0.831	0.121	0.895	95	0.717	0.074	0.925	186	0.637	0.060	0.958	224
115	1.037	0.151	0.887	94	0.982	0.098	0.865	199	0.804	0.069	0.883	270
117	0.727	0.070	0.894	216	0.712	0.048	0.881	435	0.635	0.035	0.890	656
118	0.848	0.086	0.689	195	0.521	0.035	0.932	451	0.681	0.036	0.954	710
119	0.801	0.084	0.918	181	0.629	0.048	0.939	350	0.628	0.046	0.943	375
120	0.702	0.135	0.797	54	0.684	0.108	0.793	80	1.170	0.194	0.921	73
121	0.958	0.077	0.886	309	0.724	0.040	0.912	668	0.611	0.028	0.910	968
122	0.750	0.062	0.889	296	0.759	0.042	0.941	653	0.681	0.035	0.954	770
123	0.837	0.068	0.885	301	0.677	0.034	0.949	787	0.780	0.031	0.952	1,254
124	1.331	0.086	0.929	477	1.223	0.042	0.918	1,677	1.086	0.030	0.888	2,642
125	1.362	0.081	0.881	570	1.017	0.036	0.918	1,606	0.959	0.026	0.920	2,665
126	1.129	0.085	0.877	354	1.107	0.053	0.934	883	1.135	0.041	0.973	1,543
127	1.347	0.089	0.917	458	1.291	0.045	0.937	1,680	1.189	0.032	0.950	2,729

Table B2	2-digit trade flows				3-digit trade flows				4-digit trade flows			
country	estimate	st error	fit	# obs	estimate	st error	fit	# obs	estimate	st error	fit	# obs
128	0.726	0.081	0.907	160	0.748	0.055	0.924	366	0.649	0.040	0.914	525
129	0.859	0.070	0.858	305	0.889	0.044	0.936	822	0.921	0.039	0.946	1,123
130	0.626	0.090	0.832	96	0.571	0.049	0.857	267	0.576	0.044	0.876	347
131	1.003	0.083	0.911	295	0.595	0.035	0.948	570	0.591	0.030	0.936	777
132	0.980	0.079	0.812	304	0.900	0.040	0.925	1,007	0.772	0.028	0.885	1,507
133	1.067	0.069	0.907	480	0.970	0.042	0.920	1,089	0.868	0.029	0.962	1,783
134	1.343	0.096	0.980	393	1.328	0.053	0.979	1,245	1.420	0.046	0.964	1,948
135	1.206	0.072	0.918	558	1.026	0.037	0.975	1,510	1.025	0.029	0.972	2,484
136	1.386	0.087	0.926	503	1.392	0.047	0.907	1,778	1.287	0.034	0.911	2,856
137	1.850	0.099	0.948	705	1.317	0.039	0.960	2,245	1.093	0.024	0.945	4,064
138	1.105	0.074	0.955	451	0.818	0.034	0.949	1,145	0.756	0.027	0.938	1,616
139	0.727	0.075	0.806	186	0.647	0.047	0.867	377	0.565	0.035	0.918	536
140	0.957	0.066	0.882	415	0.870	0.037	0.924	1,124	0.846	0.031	0.947	1,464
142	2.807	0.143	0.945	769	2.114	0.060	0.954	2,521	1.888	0.041	0.944	4,195
143	1.441	0.074	0.904	763	1.325	0.039	0.981	2,313	1.197	0.028	0.962	3,732
144	2.373	0.115	0.981	859	2.421	0.062	0.968	3,085	2.330	0.047	0.983	4,993
145	3.710	0.190	0.874	759	3.326	0.087	0.907	2,942	2.948	0.059	0.900	5,075
146	1.293	0.077	0.919	559	1.074	0.039	0.916	1,546	0.973	0.026	0.937	2,817
147	1.311	0.072	0.947	662	1.253	0.042	0.939	1,810	1.245	0.032	0.954	2,979
148	2.149	0.121	0.937	626	2.064	0.058	0.940	2,567	1.765	0.037	0.960	4,455
149	1.777	0.088	0.918	815	1.779	0.050	0.930	2,510	1.542	0.032	0.944	4,529
150	1.234	0.074	0.860	553	1.061	0.036	0.969	1,708	1.045	0.027	0.913	2,960
151	1.704	0.093	0.946	676	1.467	0.043	0.982	2,280	1.417	0.032	0.979	4,010
152	2.829	0.150	0.954	715	2.586	0.070	0.953	2,711	2.475	0.055	0.972	4,037
154	2.111	0.110	0.966	732	1.937	0.053	0.966	2,690	1.647	0.036	0.950	4,184
155	1.132	0.075	0.894	455	1.340	0.048	0.951	1,582	1.095	0.029	0.921	2,757
156	0.643	0.059	0.855	239	0.588	0.037	0.883	501	0.559	0.029	0.873	745
157	1.328	0.084	0.899	503	1.134	0.047	0.929	1,144	1.124	0.035	0.921	2,093
158	1.787	0.102	0.962	615	2.006	0.064	0.968	1,961	1.420	0.034	0.951	3,506
159	1.569	0.091	0.917	597	1.458	0.045	0.934	2,076	1.362	0.034	0.957	3,196

Table B2	2-digit trade flows				3-digit trade flows				4-digit trade flows			
country	estimate	st error	fit	# obs	estimate	st error	fit	# obs	estimate	st error	fit	# obs
161	1.171	0.095	0.979	302	0.962	0.049	0.976	786	0.998	0.048	0.975	875
162	1.218	0.092	0.916	349	0.952	0.042	0.922	1,030	0.757	0.028	0.928	1,497
164	0.807	0.056	0.916	416	0.840	0.038	0.932	984	0.752	0.027	0.936	1,505
165	1.266	0.085	0.985	439	1.312	0.051	0.978	1,329	1.144	0.035	0.966	2,117
166	1.845	0.105	0.957	622	1.542	0.046	0.972	2,237	1.493	0.042	0.989	2,492
167	1.158	0.116	0.970	200	1.469	0.077	0.992	731	1.205	0.057	0.984	886
168	2.033	0.103	0.941	784	1.377	0.040	0.982	2,320	1.248	0.035	0.979	2,594
169	1.470	0.089	0.964	545	1.240	0.042	0.969	1,771	1.226	0.034	0.957	2,582
170	1.306	0.096	0.982	373	1.445	0.059	0.973	1,193	1.346	0.041	0.970	2,119
172	1.981	0.101	0.942	774	1.769	0.050	0.952	2,544	1.345	0.028	0.981	4,515
173	1.403	0.100	0.905	391	1.340	0.060	0.929	1,003	1.248	0.049	0.923	1,293
174	1.070	0.074	0.798	420	0.983	0.040	0.916	1,205	0.964	0.029	0.910	2,238
175	0.970	0.062	0.827	482	0.834	0.036	0.899	1,057	0.798	0.024	0.946	2,134
176	0.624	0.073	0.654	148	0.555	0.048	0.806	263	0.506	0.042	0.806	288
177	0.661	0.061	0.903	234	0.822	0.047	0.959	611	0.711	0.034	0.953	888
178	0.785	0.070	0.853	248	0.636	0.040	0.921	495	0.598	0.034	0.923	613
179	0.884	0.087	0.795	205	0.601	0.045	0.900	353	0.469	0.033	0.839	410
180	0.773	0.070	0.761	244	0.696	0.046	0.831	460	0.590	0.036	0.912	545
all	0.955	0.006	0.968	53,503	0.928	0.003	0.991	142,231	0.869	0.003	0.991	221,430

Table B3 GI slope estimates of Rank-Size rule; BI > 1, sector dimension

sector	2-digit trade flows estimate	st error	fit	# obs	sector	2-digit trade flows estimate	st error	fit	# obs
01	1.048	0.052	0.873	797	58	2.301	0.170	0.950	365
02	1.191	0.073	0.953	532	59	2.706	0.163	0.927	551
03	0.787	0.025	0.903	2,024	61	1.064	0.046	0.925	1,078
04	1.283	0.067	0.894	737	62	2.177	0.124	0.969	615
05	0.999	0.033	0.884	1,866	63	1.366	0.058	0.946	1,109
06	0.699	0.027	0.898	1,376	64	1.294	0.086	0.946	450
07	0.646	0.022	0.764	1,710	65	1.472	0.065	0.956	1,032
08	1.103	0.046	0.926	1,138	66	1.107	0.049	0.949	1,017
09	1.334	0.061	0.978	957	67	1.696	0.089	0.970	722
11	1.294	0.064	0.905	814	68	1.052	0.049	0.944	939
12	0.904	0.038	0.967	1,122	69	2.156	0.117	0.981	678
21	0.914	0.035	0.939	1,395	71	2.889	0.208	0.980	385
22	0.796	0.036	0.891	980	72	2.581	0.182	0.983	403
23	0.752	0.045	0.831	548	73	1.968	0.146	0.918	364
24	0.926	0.038	0.901	1,177	74	3.005	0.212	0.924	401
25	1.041	0.072	0.873	413	75	2.004	0.159	0.972	319
26	0.752	0.028	0.895	1,443	76	2.020	0.140	0.904	417
27	0.813	0.033	0.954	1,213	77	2.248	0.138	0.987	531
28	0.777	0.031	0.863	1,260	78	3.113	0.254	0.925	301
29	1.092	0.038	0.960	1,695	79	1.388	0.086	0.979	522
32	0.984	0.073	0.864	366	81	2.108	0.114	0.982	687
33	1.107	0.048	0.772	1,054	82	1.770	0.098	0.941	657
34	0.908	0.064	0.886	404	83	1.259	0.069	0.909	668
35	1.022	0.069	0.968	438	84	1.193	0.047	0.877	1,304
41	0.977	0.062	0.944	494	85	1.366	0.070	0.887	758
42	0.853	0.037	0.898	1,059	87	2.435	0.173	0.966	397
43	1.256	0.074	0.972	580	88	1.461	0.106	0.911	379
51	1.890	0.124	0.989	461	89	1.915	0.101	0.960	725
52	0.966	0.049	0.967	781	93	1.170	0.059	0.965	779
53	2.057	0.117	0.985	620	94	0.797	0.031	0.954	1,345
54	1.405	0.073	0.967	744	95	1.456	0.146	0.891	199
55	1.284	0.056	0.977	1,036	97	0.804	0.050	0.876	526
56	1.124	0.054	0.954	862	99	0.808	0.041	0.800	784
all	0.955	0.006	0.968	53,503					

Table B3 sector	3-digit trade flows estimate	st error	fit	# obs	sector	3-digit trade flows estimate	st error	fit	# obs
001	0.862	0.041	0.977	880	653	1.631	0.094	0.920	607
011	1.019	0.052	0.891	772	654	0.866	0.055	0.970	501
012	0.869	0.072	0.941	295	655	1.545	0.096	0.917	517
014	1.024	0.056	0.913	680	656	1.362	0.080	0.970	576
022	1.201	0.082	0.951	433	657	1.586	0.083	0.987	727
023	0.857	0.063	0.909	375	658	1.141	0.048	0.981	1,154

Table B3	3-digit trade flows					3-digit trade flows			
sector	estimate	st error	fit	# obs	sector	estimate	st error	fit	# obs
024	1.204	0.075	0.887	521	659	0.884	0.047	0.970	708
025	1.081	0.065	0.922	560	661	1.096	0.048	0.933	1,027
034	0.736	0.028	0.938	1,399	662	1.626	0.106	0.915	471
035	0.722	0.031	0.961	1,069	663	1.921	0.121	0.990	500
036	0.730	0.024	0.913	1,801	665	2.241	0.127	0.942	625
037	0.758	0.031	0.944	1,163	666	1.684	0.099	0.947	580
041	1.211	0.104	0.861	269	667	0.800	0.041	0.878	750
042	0.606	0.033	0.816	668	671	0.737	0.039	0.954	724
043	1.238	0.095	0.883	336	672	1.189	0.067	0.977	630
044	1.061	0.075	0.904	399	673	1.519	0.085	0.969	634
045	0.878	0.064	0.981	374	674	1.983	0.129	0.913	471
046	0.987	0.061	0.951	522	676	1.219	0.089	0.968	372
047	1.139	0.066	0.958	594	677	1.711	0.104	0.977	541
048	1.487	0.070	0.981	908	678	2.395	0.160	0.975	450
054	1.045	0.039	0.937	1,452	679	1.804	0.109	0.971	549
056	1.186	0.055	0.936	932	681	1.009	0.065	0.928	475
057	0.792	0.028	0.876	1,613	682	0.700	0.042	0.923	543
058	1.102	0.044	0.936	1,247	683	0.852	0.068	0.945	314
061	0.676	0.026	0.891	1,327	684	1.061	0.056	0.931	723
062	1.413	0.065	0.963	954	685	0.777	0.048	0.909	525
071	0.623	0.025	0.777	1,230	686	0.916	0.057	0.943	516
073	0.633	0.029	0.895	959	689	0.693	0.044	0.926	501
074	0.562	0.029	0.873	750	691	2.046	0.116	0.973	624
075	0.656	0.025	0.964	1,364	692	1.647	0.078	0.995	882
081	1.107	0.046	0.928	1,147	693	1.977	0.106	0.953	696
091	1.072	0.061	0.959	624	694	2.139	0.147	0.926	421
098	1.297	0.062	0.979	883	695	1.877	0.117	0.961	518
111	1.186	0.064	0.982	686	696	1.977	0.122	0.915	528
112	1.253	0.063	0.906	792	697	1.900	0.095	0.934	795
121	0.756	0.034	0.960	1,005	699	1.375	0.081	0.955	583
122	1.132	0.058	0.970	758	711	1.801	0.133	0.971	369
211	0.926	0.036	0.930	1,351	712	1.636	0.140	0.979	273
212	0.839	0.038	0.953	956	713	2.287	0.178	0.988	329
222	0.801	0.043	0.905	683	714	1.631	0.125	0.981	340
223	0.625	0.025	0.916	1,290	716	2.710	0.163	0.972	552
232	0.664	0.042	0.767	497	718	1.355	0.097	0.971	391
233	1.588	0.109	0.957	423	721	1.789	0.117	0.985	470
244	0.659	0.076	0.765	152	722	1.959	0.151	0.957	336
245	0.784	0.042	0.922	699	723	1.784	0.116	0.993	470
246	0.870	0.055	0.964	507	724	1.575	0.138	0.921	260
247	0.642	0.030	0.868	903	725	1.381	0.119	0.897	271
248	1.118	0.050	0.891	1,013	726	1.467	0.133	0.935	243
251	1.047	0.073	0.874	412	727	1.798	0.120	0.936	452
261	0.605	0.053	0.863	264	728	2.400	0.181	0.954	353

Table B3	3-digit trade flows					3-digit trade flows			
sector	estimate	st error	fit	# obs	sector	estimate	st error	fit	# obs
263	0.645	0.027	0.889	1,153	736	1.924	0.144	0.909	357
264	0.394	0.035	0.859	248	737	1.332	0.104	0.978	329
265	0.620	0.038	0.828	535	741	2.543	0.172	0.970	438
266	1.847	0.113	0.977	538	742	2.542	0.182	0.961	391
268	0.555	0.037	0.880	457	743	2.555	0.194	0.901	348
269	1.109	0.070	0.990	502	744	1.952	0.130	0.984	452
271	0.439	0.030	0.839	436	745	1.633	0.138	0.913	281
273	1.480	0.065	0.976	1,023	749	2.921	0.199	0.941	430
274	0.852	0.067	0.921	326	751	1.869	0.146	0.921	326
277	0.641	0.038	0.960	555	752	1.767	0.143	0.971	307
278	0.994	0.044	0.909	1,008	759	1.536	0.143	0.956	232
281	0.709	0.048	0.903	442	761	1.484	0.101	0.894	428
282	1.288	0.072	0.986	646	762	1.270	0.100	0.839	324
286	0.496	0.061	0.898	132	764	2.116	0.152	0.962	386
287	0.661	0.028	0.863	1,132	771	1.953	0.118	0.972	551
288	1.264	0.056	0.989	1,032	772	1.926	0.118	0.992	531
289	0.723	0.046	0.935	490	773	1.549	0.084	0.973	676
291	1.022	0.038	0.962	1,472	774	1.673	0.142	0.973	279
292	1.054	0.038	0.962	1,507	775	2.343	0.143	0.955	534
322	0.935	0.072	0.865	340	776	1.301	0.097	0.953	361
323	0.963	0.076	0.958	325	778	2.625	0.165	0.984	504
333	1.052	0.053	0.757	776	781	2.460	0.217	0.811	258
334	1.067	0.043	0.924	1,206	782	1.977	0.158	0.949	314
335	1.012	0.053	0.971	734	783	1.374	0.088	0.974	493
341	0.909	0.064	0.885	404	784	3.420	0.277	0.906	304
351	1.026	0.069	0.967	441	785	1.472	0.113	0.946	339
411	0.983	0.062	0.943	498	786	1.823	0.113	0.984	524
423	0.971	0.064	0.978	456	791	1.359	0.094	0.965	416
424	0.788	0.034	0.890	1,085	792	1.498	0.130	0.943	267
431	1.258	0.074	0.972	578	793	0.991	0.062	0.957	517
511	1.949	0.135	0.965	417	812	2.105	0.113	0.977	688
512	1.543	0.094	0.990	539	821	1.737	0.096	0.927	658
513	1.384	0.094	0.901	432	831	1.240	0.068	0.902	669
514	1.602	0.117	0.976	373	842	1.124	0.045	0.916	1,227
515	1.282	0.087	0.954	430	843	1.177	0.048	0.927	1,191
516	1.426	0.103	0.990	381	844	0.966	0.042	0.921	1,070
522	0.907	0.046	0.962	778	845	1.219	0.054	0.934	1,031
523	1.446	0.080	0.983	651	846	1.046	0.045	0.936	1,093
524	0.690	0.065	0.916	225	847	1.223	0.056	0.946	941
533	2.052	0.117	0.983	619	848	1.058	0.051	0.931	848
541	1.357	0.070	0.980	744	851	1.332	0.068	0.881	759
551	0.829	0.039	0.980	926	871	2.023	0.167	0.973	295
553	1.425	0.076	0.945	712	872	1.646	0.119	0.975	384
554	1.342	0.065	0.977	850	874	2.068	0.151	0.981	374

Table B3	3-digit trade flows			
sector	estimate	st error	fit	# obs
562	1.124	0.054	0.952	863
582	2.151	0.178	0.950	291
583	2.291	0.157	0.943	428
584	2.120	0.150	0.982	399
591	1.576	0.094	0.942	566
592	1.312	0.077	0.972	581
598	2.247	0.156	0.963	414
611	0.919	0.043	0.920	924
612	0.933	0.046	0.962	826
613	0.956	0.058	0.914	547
621	2.360	0.147	0.960	518
625	2.030	0.116	0.946	609
628	1.279	0.084	0.929	469
634	1.134	0.052	0.945	956
635	1.350	0.061	0.974	987
641	1.072	0.090	0.898	286
642	1.595	0.075	0.957	900
651	1.346	0.062	0.962	940
652	1.306	0.056	0.963	1,072

	3-digit trade flows			
sector	estimate	st error	fit	# obs
881	1.869	0.160	0.981	273
882	2.462	0.202	0.955	298
883	0.917	0.050	0.976	671
884	1.316	0.096	0.966	376
885	0.945	0.076	0.872	311
892	1.618	0.087	0.984	696
893	2.093	0.113	0.980	690
894	1.189	0.076	0.947	492
895	2.211	0.137	0.922	522
896	1.083	0.065	0.970	561
897	1.103	0.062	0.949	637
898	1.690	0.124	0.988	373
899	1.413	0.077	0.968	675
931	1.177	0.060	0.962	776
941	0.807	0.031	0.947	1,349
951	1.457	0.146	0.891	199
971	0.802	0.049	0.877	526
999	0.808	0.041	0.799	777
all	0.928	0.003	0.991	142,231

Table B3	4-digit trade flows			
sector	estimate	st error	fit	# obs
0011	0.873	0.049	0.966	640
0012	0.559	0.032	0.923	606
0013	1.017	0.092	0.887	243
0014	0.992	0.060	0.941	554
0015	0.940	0.062	0.948	466
0019	0.485	0.034	0.969	403
0111	0.846	0.045	0.853	705
0112	0.713	0.050	0.916	402
0113	1.154	0.090	0.930	327
0114	1.190	0.085	0.842	396
0115	0.818	0.060	0.878	366
0116	0.993	0.070	0.859	400
0118	0.869	0.044	0.948	770
0121	0.799	0.078	0.915	210
0129	0.836	0.052	0.986	519
0142	1.098	0.070	0.897	497
0149	1.003	0.055	0.921	659
0223	1.507	0.114	0.908	348
0224	1.119	0.075	0.954	444
0251	1.043	0.063	0.924	541
0252	1.155	0.076	0.937	460
0341	0.775	0.032	0.960	1,172
0342	0.728	0.029	0.939	1,276

	4-digit trade flows			
sector	estimate	st error	fit	# obs
6413	1.064	0.082	0.887	337
6415	1.264	0.087	0.965	426
6416	1.128	0.066	0.941	580
6417	1.246	0.078	0.960	517
6418	1.137	0.090	0.905	321
6419	1.658	0.120	0.955	380
6421	1.107	0.051	0.946	949
6424	1.743	0.089	0.966	773
6428	1.860	0.103	0.952	654
6512	1.098	0.062	0.946	626
6514	2.122	0.125	0.947	572
6517	1.596	0.100	0.958	511
6519	1.034	0.048	0.957	926
6521	0.952	0.044	0.925	932
6522	1.356	0.064	0.955	892
6531	1.457	0.089	0.912	539
6539	1.618	0.096	0.978	565
6542	1.134	0.078	0.951	419
6549	0.693	0.048	0.962	420
6552	0.764	0.057	0.984	359
6571	1.547	0.119	0.957	338
6573	1.943	0.137	0.933	404
6575	0.913	0.041	0.961	1,012

Table B3	4-digit trade flows					4-digit trade flows			
sector	estimate	st error	fit	# obs	sector	estimate	st error	fit	# obs
0343	0.675	0.034	0.942	779	6577	1.604	0.106	0.966	457
0344	0.604	0.030	0.947	822	6581	0.783	0.038	0.982	843
0371	0.759	0.033	0.946	1,056	6583	1.054	0.052	0.865	820
0372	0.727	0.036	0.978	802	6584	1.075	0.049	0.966	979
0411	0.895	0.077	0.945	267	6589	1.364	0.065	0.968	872
0412	1.124	0.107	0.847	222	6591	1.150	0.094	0.990	297
0421	0.546	0.034	0.903	512	6592	0.869	0.047	0.967	683
0422	0.617	0.036	0.838	585	6611	1.097	0.073	0.976	451
0451	0.981	0.097	0.910	206	6612	0.900	0.038	0.907	1,102
0452	0.936	0.082	0.888	263	6613	1.018	0.078	0.814	337
0459	0.800	0.064	0.972	314	6618	1.380	0.072	0.993	734
0481	0.932	0.056	0.958	561	6623	1.354	0.108	0.938	314
0483	1.093	0.060	0.980	673	6624	1.440	0.098	0.917	436
0484	1.472	0.074	0.973	788	6631	1.462	0.108	0.968	364
0488	1.318	0.077	0.977	586	6632	1.924	0.144	0.904	358
0541	0.787	0.042	0.965	686	6633	1.646	0.107	0.996	470
0542	0.762	0.033	0.928	1,086	6638	1.870	0.121	0.975	476
0544	0.700	0.047	0.900	437	6664	1.541	0.097	0.937	501
0545	1.049	0.040	0.971	1,381	6665	1.521	0.103	0.950	433
0546	1.225	0.065	0.923	717	6666	1.250	0.082	0.934	467
0561	1.153	0.059	0.880	770	6672	0.746	0.043	0.850	591
0565	1.118	0.054	0.935	851	6674	0.784	0.041	0.928	719
0571	0.712	0.042	0.843	578	6712	0.896	0.062	0.930	412
0572	0.778	0.043	0.902	653	6713	0.788	0.065	0.972	291
0574	0.852	0.057	0.878	453	6716	0.652	0.037	0.944	623
0575	0.688	0.044	0.913	485	6724	0.940	0.068	0.963	383
0577	0.713	0.032	0.927	978	6725	1.172	0.067	0.980	606
0579	0.656	0.025	0.906	1,421	6731	1.262	0.074	0.993	578
0583	1.006	0.044	0.954	1,030	6732	1.482	0.083	0.977	639
0585	0.983	0.043	0.985	1,043	6733	1.366	0.088	0.966	483
0586	0.953	0.048	0.946	780	6748	1.192	0.146	0.960	133
0589	0.965	0.046	0.868	870	6749	1.301	0.173	0.966	113
0611	0.596	0.027	0.854	981	6781	1.407	0.110	0.946	326
0612	0.884	0.044	0.938	818	6782	1.727	0.135	0.964	329
0616	0.827	0.042	0.862	760	6783	2.024	0.127	0.970	505
0619	0.790	0.033	0.883	1,161	6785	2.653	0.166	0.979	511
0711	0.613	0.025	0.779	1,235	6793	1.446	0.112	0.977	335
0712	0.793	0.051	0.896	480	6794	1.721	0.108	0.926	511
0741	0.555	0.029	0.868	710	6811	0.864	0.052	0.928	545
0742	0.531	0.068	0.681	121	6812	0.903	0.097	0.898	172
0752	0.688	0.047	0.952	422	6821	0.604	0.039	0.907	486
0811	1.175	0.073	0.965	519	6822	1.625	0.101	0.969	516
0812	0.874	0.040	0.916	973	6831	0.762	0.062	0.923	304
0813	0.921	0.045	0.904	826	6832	1.986	0.169	0.927	276

Table B3	4-digit trade flows					4-digit trade flows			
sector	estimate	st error	fit	# obs	sector	estimate	st error	fit	# obs
0814	0.662	0.040	0.902	558	6841	0.814	0.048	0.876	568
0819	1.612	0.092	0.984	617	6842	1.550	0.086	0.975	654
1121	1.011	0.071	0.868	404	6851	0.752	0.048	0.893	496
1123	1.218	0.069	0.923	623	6852	1.248	0.087	0.969	407
1124	0.957	0.054	0.905	637	6861	0.884	0.056	0.932	494
1211	0.694	0.032	0.951	922	6863	1.056	0.073	0.952	420
1212	0.717	0.035	0.968	844	6899	0.725	0.089	0.980	134
1213	0.746	0.039	0.974	748	6931	1.696	0.099	0.955	592
1222	1.101	0.065	0.981	576	6935	1.461	0.080	0.952	675
1223	0.815	0.045	0.917	661	6951	1.142	0.059	0.970	743
2111	1.122	0.052	0.935	916	6953	1.589	0.104	0.989	463
2112	0.866	0.048	0.962	659	6954	1.707	0.120	0.965	405
2117	0.661	0.037	0.928	641	6973	1.497	0.085	0.941	613
2119	0.655	0.024	0.912	1,539	6974	1.693	0.089	0.940	730
2222	0.780	0.082	0.887	181	6991	2.004	0.140	0.898	411
2223	0.505	0.031	0.900	522	6992	1.927	0.132	0.956	423
2224	0.691	0.039	0.941	641	6996	2.129	0.115	0.956	682
2225	0.503	0.030	0.896	548	6997	1.657	0.095	0.940	607
2226	0.642	0.036	0.948	644	6998	1.851	0.110	0.992	564
2232	0.455	0.029	0.871	501	6999	1.137	0.070	0.993	525
2234	0.648	0.069	0.956	178	7133	1.699	0.139	0.884	298
2235	0.587	0.044	0.815	361	7139	2.188	0.170	0.987	332
2238	0.588	0.024	0.925	1,227	7149	1.404	0.190	0.968	109
2331	1.552	0.109	0.951	408	7188	1.483	0.106	0.980	392
2332	1.323	0.081	0.986	531	7211	1.688	0.101	0.963	560
2471	0.943	0.067	0.946	400	7212	2.030	0.155	0.932	342
2472	0.527	0.026	0.824	810	7213	0.890	0.078	0.840	260
2479	0.589	0.039	0.968	448	7219	1.673	0.115	0.976	426
2481	0.721	0.047	0.937	470	7234	1.512	0.181	0.996	140
2482	1.098	0.071	0.838	472	7243	1.802	0.141	0.932	325
2483	0.747	0.032	0.874	1,070	7247	1.413	0.130	0.931	237
2511	1.732	0.110	0.978	495	7248	1.329	0.124	0.966	231
2512	0.924	0.068	0.852	370	7251	1.061	0.092	0.944	266
2516	0.902	0.085	0.909	227	7252	1.296	0.141	0.953	169
2517	0.912	0.071	0.845	333	7259	1.403	0.116	0.935	294
2518	0.982	0.082	0.774	290	7263	1.188	0.127	0.959	175
2613	0.599	0.069	0.777	149	7269	1.161	0.120	0.969	187
2614	0.684	0.051	0.919	357	7272	2.249	0.255	0.953	155
2681	0.529	0.041	0.866	330	7281	1.733	0.151	0.903	264
2682	0.621	0.044	0.916	395	7283	1.721	0.115	0.976	449
2683	0.591	0.040	0.962	431	7284	2.426	0.192	0.910	319
2685	0.708	0.046	0.927	470	7361	2.088	0.171	0.908	298
2686	0.877	0.056	0.949	493	7369	1.748	0.146	0.986	288
2731	1.087	0.054	0.909	814	7371	1.675	0.134	0.990	311

Table B3	4-digit trade flows					4-digit trade flows			
sector	estimate	st error	fit	# obs	sector	estimate	st error	fit	# obs
2732	1.355	0.072	0.981	713	7413	1.595	0.113	0.981	399
2733	1.162	0.078	0.948	444	7414	2.447	0.179	0.908	375
2771	0.550	0.038	0.918	415	7416	2.177	0.169	0.934	331
2772	1.011	0.072	0.961	399	7431	3.109	0.427	0.934	106
2782	0.888	0.053	0.936	572	7441	1.907	0.150	0.981	323
2783	0.842	0.041	0.967	835	7442	2.411	0.165	0.979	428
2784	0.674	0.060	0.884	250	7451	1.502	0.135	0.963	249
2785	0.826	0.049	0.886	559	7452	1.818	0.157	0.827	267
2786	0.868	0.072	0.913	288	7491	2.131	0.163	0.893	341
2789	0.982	0.047	0.956	877	7492	2.409	0.171	0.950	399
2814	0.565	0.057	0.867	198	7493	2.246	0.177	0.957	323
2815	0.639	0.046	0.908	386	7499	1.771	0.123	0.984	414
2816	0.761	0.058	0.889	345	7511	1.710	0.130	0.941	346
2871	0.636	0.041	0.924	474	7512	1.668	0.138	0.899	292
2872	0.568	0.053	0.883	226	7591	1.519	0.149	0.954	209
2873	0.415	0.030	0.789	392	7641	1.949	0.137	0.978	407
2874	0.610	0.040	0.917	464	7642	1.649	0.127	0.968	335
2875	0.632	0.041	0.921	474	7649	2.010	0.150	0.967	357
2876	0.506	0.036	0.854	385	7721	1.847	0.208	0.964	158
2879	0.708	0.033	0.922	902	7722	1.460	0.166	0.941	154
2881	1.076	0.056	0.990	751	7731	1.482	0.081	0.974	668
2882	1.253	0.055	0.986	1,021	7732	1.855	0.138	0.990	362
2911	0.759	0.028	0.947	1,463	7742	1.615	0.147	0.971	240
2919	1.025	0.045	0.959	1,052	7751	1.653	0.140	0.951	280
2922	0.591	0.028	0.927	898	7752	1.512	0.087	0.952	601
2924	0.789	0.032	0.916	1,232	7754	1.086	0.127	0.831	146
2925	1.160	0.061	0.924	718	7757	1.995	0.140	0.968	408
2926	1.094	0.076	0.888	419	7758	2.120	0.137	0.958	476
2927	0.857	0.051	0.949	575	7781	1.598	0.086	0.988	686
2929	0.909	0.038	0.957	1,166	7782	1.491	0.110	0.965	368
3221	0.915	0.071	0.863	334	7783	2.024	0.157	0.948	334
3224	1.092	0.106	0.914	211	7784	1.226	0.113	0.903	234
3231	0.774	0.072	0.954	229	7788	1.675	0.118	0.984	403
3232	1.051	0.084	0.966	314	7821	1.927	0.159	0.914	295
3341	1.070	0.048	0.951	995	7822	1.360	0.128	0.976	227
3343	0.956	0.044	0.930	956	7831	1.430	0.093	0.974	470
3344	0.964	0.043	0.931	1,018	7849	3.282	0.278	0.895	278
3345	1.215	0.063	0.977	756	7852	1.481	0.117	0.955	323
3351	1.181	0.092	0.941	328	7912	0.971	0.185	0.949	55
3352	0.903	0.059	0.961	471	7925	1.111	0.229	0.920	47
3353	0.966	0.067	0.985	419	7928	1.149	0.152	0.985	114
3354	0.990	0.062	0.989	516	7932	1.141	0.147	0.917	121
3359	0.698	0.056	0.946	306	8121	2.148	0.161	0.863	357
4111	0.628	0.042	0.917	444	8122	1.312	0.067	0.990	775

Table B3	4-digit trade flows					4-digit trade flows			
sector	estimate	st error	fit	# obs	sector	estimate	st error	fit	# obs
4113	1.083	0.086	0.921	317	8124	1.798	0.113	0.949	509
4232	0.846	0.063	0.966	366	8212	1.444	0.143	0.980	203
4239	1.502	0.118	0.935	322	8219	1.430	0.130	0.982	243
4241	0.549	0.058	0.860	180	8421	1.074	0.047	0.938	1,040
4243	0.542	0.037	0.828	422	8422	0.969	0.044	0.953	950
4249	0.771	0.036	0.905	931	8423	0.943	0.040	0.905	1,124
4313	1.187	0.083	0.962	407	8429	1.111	0.047	0.922	1,123
4314	0.677	0.036	0.918	689	8431	1.113	0.050	0.922	1,005
5111	1.668	0.247	0.937	91	8432	1.092	0.050	0.955	964
5112	1.668	0.246	0.938	92	8433	1.061	0.047	0.960	1,022
5119	1.668	0.247	0.937	91	8434	1.078	0.047	0.915	1,034
5121	1.407	0.084	0.991	559	8441	0.939	0.042	0.917	1,017
5123	1.600	0.125	0.960	326	8442	0.942	0.044	0.895	911
5157	0.696	0.072	0.984	188	8451	1.127	0.053	0.960	910
5161	1.585	0.143	0.982	246	8452	1.112	0.054	0.926	839
5162	1.683	0.118	0.974	408	8459	1.160	0.053	0.907	958
5169	0.960	0.065	0.980	437	8461	1.046	0.046	0.925	1,050
5221	1.593	0.099	0.951	516	8463	0.625	0.046	0.963	371
5222	0.718	0.046	0.937	482	8465	0.841	0.043	0.962	774
5224	1.533	0.113	0.988	365	8471	1.180	0.059	0.960	794
5225	0.804	0.044	0.962	679	8472	1.128	0.055	0.944	832
5231	1.530	0.086	0.972	626	8481	0.953	0.051	0.920	710
5239	1.082	0.071	0.962	465	8482	1.148	0.075	0.934	473
5331	1.651	0.106	0.974	487	8483	0.878	0.050	0.944	626
5334	1.905	0.102	0.973	704	8484	1.265	0.072	0.934	616
5411	1.185	0.101	0.962	277	8741	1.293	0.078	0.987	556
5413	1.398	0.091	0.983	475	8745	1.611	0.120	0.975	360
5414	1.037	0.063	0.947	536	8748	2.058	0.164	0.982	316
5415	0.916	0.068	0.969	365	8749	1.637	0.127	0.985	335
5416	1.009	0.068	0.981	441	8811	1.449	0.127	0.895	260
5417	1.466	0.079	0.977	695	8813	1.619	0.170	0.954	181
5419	1.702	0.103	0.989	549	8822	2.524	0.210	0.966	288
5513	0.826	0.039	0.980	912	8841	1.705	0.152	0.925	252
5542	1.298	0.063	0.977	836	8842	1.070	0.078	0.968	376
5543	1.386	0.080	0.981	598	8851	0.863	0.071	0.856	292
5621	0.967	0.050	0.965	740	8852	1.421	0.108	0.897	347
5622	0.628	0.041	0.931	473	8921	1.532	0.090	0.987	573
5623	0.905	0.086	0.933	224	8922	2.082	0.134	0.949	486
5629	1.119	0.066	0.953	575	8925	0.808	0.079	0.920	208
5921	1.172	0.076	0.940	480	8928	1.248	0.069	0.969	649
5922	1.256	0.076	0.971	546	8931	1.748	0.087	0.975	806
5981	0.931	0.063	0.936	443	8939	1.984	0.114	0.982	602
5989	2.307	0.172	0.972	359	8942	1.156	0.083	0.931	388
6113	0.790	0.102	0.910	120	8946	1.366	0.092	0.975	439

Table B3	4-digit trade flows					4-digit trade flows			
sector	estimate	st error	fit	# obs	sector	estimate	st error	fit	# obs
6114	0.828	0.077	0.961	232	8947	0.919	0.060	0.959	464
6282	1.917	0.121	0.973	505	8981	1.562	0.123	0.957	325
6289	1.102	0.077	0.925	415	8982	1.761	0.154	0.953	263
6341	0.802	0.037	0.939	925	8983	1.631	0.120	0.991	371
6342	1.000	0.054	0.942	697	8991	0.849	0.049	0.965	591
6343	1.255	0.075	0.959	566	8996	1.320	0.110	0.903	290
6351	1.067	0.057	0.939	711	8997	1.156	0.054	0.950	920
6353	1.199	0.062	0.932	746	8998	1.798	0.109	0.970	547
6359	1.210	0.056	0.980	921	8999	0.966	0.070	0.891	386
6411	1.085	0.107	0.751	207	9999	0.800	0.040	0.803	783
6412	1.079	0.076	0.932	400	all	0.869	0.003	0.991	221,430