CESIS Electronic Working Paper Series

Paper No. 379

A note on the granular nature of imports in German manufacturing industries

Joachim Wagner

October, 2014

The Royal Institute of technology
Centre of Excellence for Science and Innovation Studies (CESIS)
http://www.cesis.se

A note on the granular nature of imports in German manufacturing industries

Joachim Wagner

Leuphana University Lueneburg, Germany, and CESIS, KTH, Stockholm, Sweden

wagner@leuphana.de

[This version: October 9, 2014]

Abstract: This paper uses an approach recently suggested by Gabaix (Eonometrica 2011) to investigate for the first time the role of idiosyncratic shocks to the largest firms in the dynamics of imports by firms from manufacturing industries. For Germany we find evidence that imports are power-law distributed and that the distribution of imports in the industries can be characterised as fat-tailed. Results show that idiosyncratic shocks to very large firms are important for the import dynamics in 2010/2011 but not in 2009/2010.

Keywords: Imports, power law, granular residual, Germany

JEL classification: F14, E32, L60

All computations were done inside the Research Data Centre of the Statistical Office of Berlin-Brandenburg. The data used are confidential but not exclusive; see www.forschungsdatenzentrum.de for information how to access the data. To facilitate replication the Stata do-files used are available from the author on request.

1. Motivation

Exports and imports tend to be highly concentrated. There is empirical evidence for a number of countries that a small share of firms are responsible for the lion's share of international trade (see World Trade Organization (2008), p. 54). Bernard et al. (2011, p.10) call this one of the striking features of international trade data. Germany, one of the leading actors on the world market for goods and services, is a case in point. More than half of total exports and imports in Germany stem from the largest 50 trading firms (see Wagner (2012a) for details).

Detailed analyses that use firm-level data to decompose the overall change in exports in Germany during the "Great Export Collapse" in 2008/09 and during the "Great Export Recovery" in 2009/2010 reveal that a small fraction of firms from the largest size class is responsible for shaping the big picture that is familiar from published aggregate figures. This illustrates that a macroeconomic development – the change in total exports from year to year – can be driven by a small number of large firms (see Wagner (2013a, 2013b)). This is in line with a view recently presented by Gabiax (2011) who argues that many economic fluctuations are attributable to the incompressible "grains" of economic activity, the large firms. Therefore, he names this view the "granular" hypothesis.

We have evidence that supports this granular view for the German manufacturing sector, where idiosyncratic shocks in the largest firms are important for an understanding of aggregate volatility (see Wagner (2012b)), and similar evidence for manufacturing exports from Germany ((see Wagner (2013a, 2013b)). This paper contributes to the literature by investigating the granular nature of imports for the first time. To anticipate the most important findings, we find evidence for a granular nature of imports in manufacturing industries in Germany for 2010/2011 but not for 2009/2010.

The rest of the paper is organized as follows: Section 2 introduces the enterprise level data used in this study. Section 3 presents the empirical approach applied to investigate the role of idiosyncratic shocks to the largest firms for the overall change in imports and discusses the results for Germany. Section 4 concludes.

2. Data

Information on imports is available from the statistic on foreign trade (*Außenhandelsstatistik*). This statistic is based on two sources. One source is the reports by German firms on transactions with firms from countries that are members of the European Union (EU); these reports are used to compile the so-called *Intrahandelsstatistik* on intra-EU trade. The other source is transaction-level data collected by the customs on trade with countries outside the EU (the so-called *Extrahandelsstatistik*).[1] Data in the statistic of foreign trade are transaction-level data, i.e. they relate to one transaction of a German firm with a firm located outside Germany at a time. Published data from this statistic report imports aggregated at the level of goods imported and by country of origin.

For the reporting years 2009 to 2011 these transaction-level data have been aggregated at the level of the importing firm for the first time. For each importing firm that reported either to the statistic on intra-EU trade, or to the statistic on trade with countries outside the EU, we know from these data the value of imports.

Using the firms' registration number for turnover tax statistics these data were matched with the enterprise register system (*Unternehmensregister-System*). For

[1] Note that firms with a value of imports from EU-countries that does not exceed 400,000 Euro in 2009 do not have to report to the statistic on intra-EU trade. For trade with firms from non-member countries all transactions that exceed 1,000 Euro are registered. For details see Statistisches Bundesamt, Qualitätsbericht Außenhandel, Januar 2011.

enterprises from manufacturing industries this matching made it possible to add information (that is taken from a regular survey of manufacturing firms) on industry affiliation.

These newly available data on import activities of firms from German manufacturing industries are used in this paper to look at the role of idiosyncratic shocks to large firms in shaping the dynamics of imports from 2009 to 2010, and from 2010 to 2011, respectively.

3. Empirical investigation

Standard macroeconomic reasoning usually discards the possibility that idiosyncratic microeconomic shocks to firms may lead to large aggregate fluctuations by referring to a diversification argument.[2] A classical case in point is the argument put forward by Lucas (1977) that such microeconomic shocks would average out and, therefore, should only have negligible aggregate effects. In a recent Econometrica paper Gabaix (2011) proposes that, contrary to this traditional view, idiosyncratic firm-level shocks can indeed explain an important part of aggregate economic movements and provide a micro-foundation for aggregate shocks. He shows that the “averaging out” argument breaks down if the size distribution of firms is fat-tailed and very large firms play an important role in an economy. This is the case in the United States, where, according to the findings of Gabaix (2011), the idiosyncratic movements of the largest 100 firms appear to explain about one-third of variations in output growth. Wagner (2012b, 2013a, 2013b) reports similar evidence for the manufacturing sector in Germany and for German exports of goods. He finds that idiosyncratic shocks in

[2] This section builds on the investigation of the granular nature of the German manufacturing sector in Wagner (2012b).

the largest firms are important for an understanding of aggregate volatility in German manufacturing industries and its exports.

Gabaix (2011) argues that many economic fluctuations are attributable to the incompressible "grains" of economic activity, the large firms. Therefore, he names this view the "granular" hypothesis. The granular view does not neglect the role of aggregate shocks like changes in monetary, fiscal, and exchange rate policy as important drivers of macroeconomic activity. It only argues that such aggregate shocks are not the only important drivers, and that firm specific idiosyncratic shocks, too, are an important, and possibly the major, part of the origin of business-cycle fluctuations (Gabaix 2011, p. 764).

As said the "averaging out" argument of standard macroeconomic reasoning breaks down if the size distribution of firms is fat-tailed and very large firms play an important role in an economy. From the percentage shares of the largest enterprises in total imports in manufacturing industries in West Germany[3] in 2010 and 2011 that are documented in Table 1 it is evident that the imports of manufacturing enterprises are highly concentrated. The very large firms, therefore, represent a large part of the import activity in the manufacturing sector.

[Table 1 near here]

In Table 2 and Table 3 the estimated power law exponents for imports are reported for all firms and for firms from 24 manufacturing industries.[4] A power law is a relation of the type $Y = k*X^{ß}$, where Y and X are variables of interest, ß is the power

[3] This study looks at West Germany only. A separate analysis of the imports from the East German manufacturing sector is not possible because the number of firms in many industries is far too small.

[4] The industries are at the 2-digit level. For a definition of industries see the appendix table.

law exponent, and k is a constant.[5] A popular way to estimate the power law exponent ß for the firm size distribution (where firm size is measured by exports here) is to compute the rank of each firm in the size distribution and to run an OLS regression of log(rank) on a constant and log(size). The estimated regression coefficient of log(size) is an estimate for ß. Gabaix and Ibragimov (2011) show that this procedure leads to strongly biased estimates in small samples. They provide a simple practical remedy for this bias by suggesting to use rank – ½ instead of rank and then run log(rank – ½) = *k* - ß*log(size). They show that the shift of ½ is optimal and reduces the bias to a leading order. Note that the standard error of ß is not the OLS standard error reported by the computer program, but is asymptotically given by $(2/n)^{½}$ *|b| (where n is the number of firms used in the estimation).

The estimated power-law coefficient for exports is statistically significantly different from zero at an error level of less than 1 percent for all imports and for imports in every industry (except industry 12, where the number of importing firms is tiny). According to the R^2-value the fit is rather tight. These results indicate that imports are power-law distributed in all industries. Descriptive results, therefore, indicate that the distribution of imports from the German manufacturing sector as a whole and from the various industries that are part of it can be characterised as fat-tailed.

[Table 2 and Table 3 near here]

To test for the granular nature of imports from German manufacturing industries the data for enterprises from 22 of the 24 manufacturing industries that are

[5] Gabaix (2009) is a comprehensive survey of power laws and applications in economics and finance.

described above are used and the role of the 10 largest firms in each industry is considered.[6] The empirical approach closely follows Gabaix (2011, p. 750ff.). The idiosyncratic firm-level sales shock is measured by the "granular residual" that is computed as follows. g_{it} is the growth rate of imports for firm i and year t, computed as log($imports_{it}$) – log($imports_{it-1}$). $g10_t$ is the average of the growth rates of the 10 largest firms (according to imports in year t-1) in an industry. The granular residual is a weighted sum of the 10 largest firm's growth rate minus $g10_t$, where the weights are the shares of the firms in total imports of all firms in an industry in year t-1. Here, t refers to 2010 (2011) and t-1 refers to 2009 (2010).

The growth rate of total imports in an industry, defined as log(total imports in t) minus log(total imports in t-1), is regressed on the granular residual from the industry using Ordinary Least Squares (OLS). Results are reported in the first column of Table 4. For 2009/2010 these results are not supportive of the granular hypothesis. The estimated coefficient for the granular residual is not statistically significant. If only aggregate shocks were important for the growth rate of total imports in an industry, then the R^2 of the regressions in Table 4 would be zero. It is. Idiosyncratic movements of the top 10 firms in an industry cannot explain a large fraction of import fluctuations. Results for 2010/2011 differ considerably. The estimated coefficient for the granular residual is statistically significant, and the R^2 is different from zero. Here, we have empirical evidence that idiosyncratic movements of the top 10 firms in an industry can explain a fraction of import fluctuations.

[Table 4 near here]

[6] Industry 12 (manufacture of tobacco products) and industry 19 (manufacture of coke and refined petroleum products) were dropped due to the small number of firms from these industries.

However, it is well known that results estimated by OLS can be highly sensitive to a small fraction of observations that lay far away from the majority of observations in the sample. As a robustness check, therefore, we investigate whether the results reported depend on extreme observations, or outliers. Rousseeuw and Leroy (1987) distinguish three types of outliers that influence the OLS estimator: vertical outliers, bad leverage points, and good leverage points. Verardi and Croux (2009, p. 440) illustrate this terminology in a simple linear regression framework that is used here (the generalization to higher dimensions is straightforward) as follows: "Vertical outliers are those observations that have outlying values for the corresponding error term (the *y* dimension) but are not outlying in the space of explanatory variables (the *x* dimension). Their presence affects the OLS estimation and, in particular, the estimated intercept. Good leverage points are observations that are outlying in the space of explanatory variables but that are located close to the regression line. Their presence does not affect the OLS estimation, but it affects statistical inference because they do deflate the estimated standard errors. Finally, bad leverage points are observations that are both outlying in the space of explanatory variables and located far from the true regression line. Their presence significantly affects the OLS estimation of both the intercept and the slope."

Using this terminology one can state that the popular median regression estimator (also known as Least Absolute Deviations or LAD) protects against vertical outliers but not against bad leverage points (Verardi and Croux 2009, p. 441). Full robustness can be achieved by using the so-called S-estimator that can resist contamination of the data set of up to 50% of outliers (i.e., that has a breakdown point[7] of 50 % compared to zero percent for OLS). A discussion of any details of this estimator is beyond the scope of this paper (see Verardi and McCathie (2012) for this estimator and for Stata commands to compute it).

Results computed by the S-estimator are reported in the second column of Table 4. For 2009/2010 the robust estimator identifies four outliers. These outliers are the observations from the industries 11 (beverages), 21 (basic pharmaceutical products), 22 (rubber ans plastic products) and 30 (other transport equipments). Dropping these outliers from the estimation sample does not change the big picture at all. Idiosyncratic movements of the top 10 firms in an industry cannot explain a large fraction of import fluctuations in 2009/2010.

For 2010/2011 the robust estimator identifies five outliers. These outliers are the observations from the industries 10 (food products), 21 (basic pharmaceutical products), 30 (other transport equipments), 31 (furniture) and 32 (other manufacturing). Again, dropping these outliers from the estimation sample does not change the big picture from the OLS regression, but this big picture is different from the one found for 2009/2010. Idiosyncratic movements of the top 10 firms in an industry can explain a fraction of import fluctuations in 2010/2011, and according to the R^2-value this fraction is considerably larger if the outliers are dropped from the estimation sample.

4. Discussion

The bottom line, then, is that we find evidence for a granular nature of imports in manufacturing industries in Germany for 2010/2011 but not for 2009/2010. The reasons for this difference between the two periods under investigation are not at all

[7] The breakdown point of an estimator is the highest fraction of outliers that an estimator can withstand, and it is a popular measure of robustness.

obvious. According to the results reported in Table 2 and Table 3 imports are power-law distributed in all industries both in 2010 and 1011, and the distribution of imports in the industries can be characterised as fat-tailed. This seems to be a structural characteristic of the German manufacturing sector that does not vary (at least not in the short run). The different results with regard to the role of idiosyncratic shocks to the largest firms for the two period might be due to the fact that imports of the firms included in the sample used here increased by 32.4 percent during the "Great Import Recovery" in 2009/2010 (that followed after the large drop in imports during the Great Recession of 2008/2009), but grew only more moderately by 16.9 percent during 2010/2011. Maybe, the huge import increase in 2009/2010 was so large that even large shocks to large firms had no decisive influence on import dynamics as a whole. Given that data for imports at the firm-level are available in Germany for 2009 to 2011 as yet, future research using data for later years with smaller overall growth rates of imports might shed more light on this.

The different findings for the two periods investigated in this paper illustrate that empirical results based on data for one period cannot be a sound basis for any conclusions regarding the validity of theoretical reasoning. Stylized facts that can guide economic theory and policy should be based on empirical investigations that replicate results for different data sets from many different years. Therefore, the jury is still out with respect to a decision on the granular nature of imports in German manufacturing industries. That said, the granular approach recently introduced by Gabaix (2011) offers a highly useful tool for the analysis of import dynamics, too.

References

Bernard, A. B., J. B. Jensen, S. J. Redding and P. K. Schott (2011): The Empirics of Firm Heterogeneity and International Trade. National Bureau of Economic Research NBER Working Paper 17672.

Gabaix, X. (2009): Power Laws in Economics and Finance, *Annual Review of Economics* 1, 255-293.

Gabaix, X. (2011): The Granular Origins of Aggregate Fluctuations, *Econometrica* 79, 733-772.

Gabaix, X. and R.Ibragimov (2011): Rank -1/2: A Simple way to improve the OLS estimation of tail exponents, *Journal of Economics and Business Statistics* 29, 24-39.

Lucas, R. E. (1977): Understanding Business Cycles, *Carnegie-Rochester Conference Series on Public Policy* 5, 7-29.

Rousseeuw, P. J. and A. M. Leroy (1987): Robust Regression and Outlier Detection. John Wiley and Sons, New York.

Verardi, V. and C. Croux (2009): Robust regression in Stata, *The Stata Journal* 9, 439-453.

Verardi, V. and A. McCathie (2012): The S-estimator of multivariate location and scatter in Stata, *The Stata Journal* 12, 299-307.

Wagner, J. (2012a): Trading many goods with many countries: Exporters and importers from German manufacturing industries. *Review of Economics* 63, 170-186.

Wagner, J. (2012b): The German Manufacturing Sector is a Granular Economy, Applied Economics Letters 19, 1663-1665.

Wagner, J. (2013a): The Granular Nature of the Great Export Collapse in German Manufacturing Industries, 2008/2009, Economics: The Open-Access, Open-Assessment E-Journal, Vol. 7, 2013-5.

Wagner, J. (2013b): The Great Export Recovery in German Manufacturing Industries, 2009/2010. *Review of Economics* 64, 325-340.

World Trade Organization (2008): World Trade Report 2008. Trade in a Globalizing World. Geneva: WTO Publications.

Table 1: Concentration of imports in enterprises from German manufacturing industries, 2010 and 2011

	Share of largest … importers in total imports (percent)			
	10	50	100	Number of importers
2010	39.05	59.22	65.82	10,870
2011	41.92	58.87	65.23	11,064

Table 2: Estimated power law exponents for imports in manufacturing industries, West Germany, 2010

Industry	ß	t-value	R^2	Number of enterprises
All	-0.308	-73.72	0.706	10,870
10	-0.446	-18.80	0.801	707
11	-0.312	-7.55	0.818	114
12	-0.393	-1.41	0.878	8
13	-0.256	-12.14	0.688	295
14	-0.394	-9.08	0.713	165
15	-0.364	-5.61	0.766	63
16	-0.245	-11.14	0.700	248
17	-0.313	-13.17	0.699	347
18	-0.188	-10.48	0.689	220
19	-0.653	-3.32	0.888	22
20	-0.383	-16.49	0.756	544
21	-0.373	-7.14	0.685	102
22	-0.283	-21.49	0.718	924
23	-0.297	-14.09	0.701	397
24	-0.377	-13.40	0.733	359
25	-0.235	-27.96	0.692	1,564
26	-0.342	-18.73	0.730	702
27	-0.322	-19.57	0.720	766
28	-0.264	-31.75	0.690	2,016
29	-0.435	-12.94	0.736	335
30	-0.415	-6.60	0.767	87
31	-0.262	-12.10	0.734	293
32	-0.284	-13.78	0.711	380
33	-0.284	-10.30	0.652	212

Note: For a definition of the industries see the appendix table. The power law exponent ß and its standard error are estimated by the method suggested in Gabaix and Ibragimov (2011); see text.

Table 3: Estimated power law exponents for imports in manufacturing industries, West Germany, 2010

Industry	ß	t-value	R^2	Number of enterprises
All	-0.230	-74.38	0.697	11,064
10	-0.425	-18.69	0.774	699
11	-0.254	-7.42	0.623	110
12	-0.384	-1.87	0.884	7
13	-0.272	-11.96	0.721	286
14	-0.464	-9.11	0.810	166
15	-0.441	-5.24	0.814	55
16	-0.241	-11.20	0.724	251
17	-0.299	-13.19	0.685	348
18	-0.168	-10.86	0.691	236
19	-0.564	-3.00	0.808	18
20	-0.353	-16.43	0.720	540
21	-0.462	-7.25	0.776	105
22	-0.276	-22.12	0.717	979
23	-0.286	-14.21	0.707	404
24	-0.403	-13.27	0.743	352
25	-0.227	-28.55	0.688	1,630
26	-0.337	-18.55	0.773	688
27	-0.304	-19.84	0.707	787
28	-0.265	-31.98	0.678	2,046
29	-0.409	-13.21	0.725	349
30	-0.459	-6.67	0.817	89
31	-0.249	-12.23	0.715	299
32	-0.297	-13.58	0.729	369
33	-0.268	-11.20	0.621	251

Note: For a definition of the industries see the appendix table. The power law exponent ß and its standard error are estimated by the method suggested in Gabaix and Ibragimov (2011); see text.

Table 4: Explanatory power of the granular residual for import growth in manufacturing industries, West Germany, 2009/2010 and 2010/2011

Dependent variable: import growth 2009/2010 (percentage)

Estimation method:		OLS	S-estimator
Granular residual 2009/2010	ß	-0.0015	0.0014
	p	0.667	0.854
Constant	ß	17.240	19.211
	P	0.000	0.000
Number of industries		22	18
R^2		0.010	0.0022

Dependent variable: import growth 2010/2011 (percentage)

Estimation method:		OLS	S-estimator
Granular residual 2010/2011	ß	-0.0062	-0.0444
	p	0.008	0.004
Constant	ß	15.189	13.943
	p	0.000	0.000
Number of industries		22	17
R^2		0.215	0.431

Note: ß is the estimated regression coefficient, p is the prob-value. For a definition of the industries see the appendix table. For a definition of the granular residual see text.

Appendix: Definition of manufacturing industries

No.	Industry
10	Manufacture of food products
11	Manufacture of beverages
12	Manufacture of tobacco products
13	Manufacture of textiles
14	Manufacture of wearing apparel
15	Manufacture of leather and related products
16	Manufacture of wood and products of wood, except furniture
17	Manufacture of paper and paper products
18	Printing and reproduction of recorded media
19	Manufacture of coke and refined petroleum products
20	Manufacture of chemicals and chemical products
21	Manufacture of basic pharmaceutical products and pharmaceutical preparations
22	Manufacture of rubber and plastic products
23	Manufacture of other non-metallic mineral products
24	Manufacture of basic metals
25	Manufacture of fabricated metal products, except machinery and equipment
26	Manufacture of computer, electronic and optical products
27	Manufacture of electrical equipment
28	Manufacture of machinery and equipment n.e.c.
29	Manufacture of motor vehicles, trailers and semi-trailers
30	Manufacture of other transport equipment
31	Manufacture of furniture
32	Other manufacturing
33	Repair and installation of machinery and equipment

Note: The 2-digit-industries are defined according to the German classification WZ 2008.