Volatility in the Small and in the Large: Diversification in Trade Networks*

Francis Kramarz[†] Julien Martin[‡] Isabelle Mejean[§]

December 2015

Abstract

This paper develops a framework to study how different sources of fluctuations and the micro-structure of trade networks shape the volatility of exports at the firm-level and in the aggregate. We consider four orthogonal shocks affecting exporter-importer trade networks – a macroeconomic shock and three individual shocks hitting respectively the exporter, its foreign partner, and the match they form. We use our framework and new data on networks connecting French exporters to European buyers over the 1995-2007 period to structurally estimate these shocks. Individual shocks are found to be a major source of fluctuations, in disaggregated as in aggregate data. Customer-related shocks matters for the volatility in the small. Firms' exposure to such shocks varies depending on the structure of their portfolio of customers: More diversified firms are better hedged against customer-related shocks. This diversification explains a sizable fraction of the dispersion of volatility across firms. In the aggregate, the relative prevalence of different types of individual shocks also varies across countries. We show that it depends on the shape of the sales distribution across sellers, buyers and seller-buyer pairs. Differences in the structure of trade networks partly explains the differences in the origins and the magnitude of the volatility of French exports across destinations.

*This paper has benefited from comments of many seminar and conference participants at Brown University, CREST-LMA, Duke University, SED Toronto, Université Laval, Université Libre de Bruxelles, Aix-Marseille School of Economics, University of Zurich, Yale University, the CEPR-ERWIT conference, Budapest-CEU, the NBER-IFM Summer Institute. We are thanksful to Ezra Oberfield for a very helpful discussion. Our work is supported by a public grant overseen by the French National Research Agency (ANR) as part of the "Investissements d'Avenir" program (Idex Grant Agreement No. ANR-11-IDEX-0003-02 / Labex ECODEC No. ANR-11-LABEX-0047). Martin thanks FRQSC for financial support (grant 2015-NP-182781).

[†]CREST-ENSAE. francis.kramarz@ensae.fr.

[‡]Université du Québec à Montréal. martin.julien@uqam.ca

[§]Ecole Polytechnique. isabelle.mejean@polytechnique.edu

1 Introduction

What explains the differences in the volatility of sales and output, across firms and across markets? The two questions have been extensively studied - most often separately - by researchers from various fields within economics.[1] This paper analyzes how the different sources of shocks hitting individual firms together with the micro-structure of the economy shapes the volatility of sales, at the firm-level as well as in the aggregate.

At the level of individual firms, volatility called in what follows *in the small* has been shown to be large, irrespective of the measure of performance used.[2] Not only is firm-level volatility high on average, it is also strongly heterogeneous across firms (Decker et al., 2014; Fort et al., 2013). While most papers rely on idiosyncratic supply shocks when explaining such dynamics, recent contributions have pointed out the role of customer-related shocks (Foster et al., 2008, 2012; Arkolakis, 2011; Kelly et al., 2013). Our paper proposes an empirical strategy which exploits highly disaggregated data to isolate different sources of shocks to growth, including idiosyncratic supply shocks as well as customer-related growth components. We quantify their relative contributions in explaining the volatility of individual sales and its heterogeneity across firms.

Individual shocks to firms are also central in the recent literature on aggregate "granular" fluctuations. When the distribution of firms' size is fat-tailed (Gabaix, 2011) and when firm-to-firm linkages are sufficiently concentrated (Acemoglu et al., 2012), idiosyncratic shocks can have a substantial effect on macroeconomic outcomes; volatility called *in the large* in the following. Here as well, the literature has mostly focused on idiosyncratic supply shocks as a source of granular fluctuations. We expand this framework and incorporate the different sources of volatility as measured at a disaggregated level. In this framework, not only the structure of economic networks but also the prevalence of different types of shocks determine the magnitude of aggregate fluctuations.

This paper more specifically focuses on shocks related to individual customers. The presence of such shocks in our statistical model is important for at least two reasons. At the level of individual producers, introducing idiosyncratic shocks related to customers means that the variance of a firm's output is affected by the structure of its portfolio of clients. Firms with a more "diversified" portfolio end up better hedged against idiosyncratic shocks hitting their customers. We provide evidence that this is indeed the case in our data. This feature also helps explain the dispersion of firms' volatility. Better diversified

[1] The volatility of individual firms has been studied in the corporate finance literature as in Thesmar and Thoenig (2011) and in industrial organization (e.g. Asker et al. (2014)). The volatility of aggregate GDPs is a classical macroeconomic question. Koren and Tenreyro (2007) and Caselli et al. (2015) thus study the difference in volatilities across countries of different levels of development. A few papers have tackled both issues together - eg. Comin and Mulani (2006) have documented the different trends in firm-level and aggregate volatility.

[2] Comin and Philippon (2006) document the increase in firm volatility using real measures (sales, employment and capital expenditures) as well as financial data (equity returns). See also Comin and Mulani (2006) on sales data, Thesmar and Thoenig (2011) based on sales and employment, Campbell (2001) using stock returns.

firms display significantly less volatile sales. The presence of customer-related shocks also matters at the aggregate level. Here, the argument is very similar to that of Gabaix (2011). Namely, if there are idiosyncratic shocks to the demand-side of the economy, shocks to the largest or the most connected buyers will show up in the aggregate. Similarly, if seller-buyer relationships are highly concentrated, shocks to the upper tail of the distribution of transactions are likely to generate granular fluctuations. From this it follows that not only the degree of sales concentration across producers but also the distribution of purchases across buyers and the distribution of sales across seller-buyer pairs matter for determining the amount of "granular" volatility. Of course, which one matters the most depends on the relative prevalence of idiosyncratic shocks to sellers, buyers, and their match, an empirical question addressed in the following.

In our empirical contribution, we consider one specific type of economic networks, namely firm-to-firm trade linkages in international markets, and the volatility of one specific outcome, the growth of exports between French firms and their EU partners. We exploit newly available data which tracks the precise structure of these transactions. The explicit identification of firms on both sides of the border constitutes one originality of the data. The data set can be viewed as a bipartite graph with sellers on one side (French exporters) and buyers on the other side (foreign importers). In this graph, the set of relationships is exhaustively observed over several years which makes the data set a good source for studying the origin of shocks in firm-to-firm trade networks. Taking the structure of the network as given, we estimate the drivers of growth at the most disaggregated level and study how, when combined, they determine the level of volatility, in the small and in the large.

Our strategy for identifying the drivers of disaggregated growth is based on a partial equilibrium model of the demand for imports, which we estimate structurally using a methodology inspired from the labor literature.[3] The demand for imports addressed by a foreign buyer to a French exporter is assumed to be affected by several aggregate and individual shocks, hitting the supply side of the economy – French exporters either collectively or individually–, the demand-side – foreign buyers again collectively or individually–, or the match formed by a seller and a buyer. The model is used to define a set of exogeneity conditions which the estimation strategy can later exploit to estimate the different growth components. In the data, "aggregate" shocks are identified as the common component of growth across all firms selling and/or buying the same type of goods in a given period. "Firm-specific" shocks are in turn measured as the residual growth component that is not explained by aggregate components. Thanks to the network structure of the data and using a specification derived from our structural model, we show how to separately identify the source of growth which is specific to a seller, and thus affects sales to all this seller's buyers, from that specific to a buyer, affecting her demand to all her sellers. Finally,

[3]Longitudinal employer-employee data have a similar graph structure as that of trade networks. This structure has been exploited to study the determinants of the wage distribution. See, among many others, Abowd et al. (1999) and Abowd et al. (2002). We apply a similar methodology to estimate the sources of trade growth in firm-to-firm trade data.

the residual of this structural decomposition is the match-specific component of growth. Hence, it becomes possible to write the yearly growth rates that we observe at the seller-buyer level into a combination of four different "shocks".

We next use the estimates to discuss the sources of volatility in the data. We first aggregate across customers within a producer's portfolio and compute the volatility *in the small*, defined as the variance of individual destination-specific sales. Second, we aggregate sales across producers to obtain a measure of volatility *in the large*, the variance of aggregate sales to a destination.[4] We then decompose these variances using our estimated shocks. In the small, aggregate shocks mechanically contribute little to explaining the variance of sales and its heterogeneity across firms. The most important drivers of the dispersion in firm-level volatilities are the seller-specific and the match-specific shocks, which respectively represent 40 and 43% of the cross-sectional variance. The sizable contribution of match-specific shocks contrasts with the related literature which, almost exclusively, focuses on supply-side drivers of fluctuations. The prevalence of buyer-related shocks also contributes in explaining the high degree of heterogeneity in firms' volatility. Namely, firms that display more diversification in sales, i.e. firms with more customers in their portfolio and a better balance of sales across buyers, have sales that are significantly less volatile. Even though this result is almost mechanical in presence of buyer-specific shocks, absent any buyer-specific source of volatility, this finding would be very difficult to explain.

Having identified the sources of volatility in the small, we next aggregate up to the country level and discuss the origin of fluctuations in the large. Aggregate shocks naturally contribute more to the volatility of a country's total sales. However, because trade data are extremely concentrated along the firm-dimension, we also expect the firm-specific components to matter. This is indeed the case. Namely, more than 80% of the volatility in aggregate sales to a destination is attributable to the combined effect of the seller-specific, the buyer-specific and the match-specific shocks. Moreover, these microeconomic shocks explain most of the differences across destinations in the magnitude of fluctuations. Here as well, shocks affecting individual sellers matter a lot, representing around 35% of aggregate fluctuations. But the most important source of fluctuations remains the seller-buyer (match) component. The magnitude of this contribution is surprising since, in the data, large firms tend to have more customers in their

[4]While performing the aggregation, the question of extensive adjustments naturally springs up. In the literature on firms' dynamics, the entry and exit of firms into a market is a key driver of fluctuations (see, for instance, Hopenhayn (1992) or Bilbiie et al. (2012)). This is potentially the case in our framework, as well. At the aggregate level, adjustments in the number of firms serving a given market contribute to the dynamic of output (Lee and Mukoyama, 2015). Less discussed in the literature is the fact that extensive adjustments can also occur *within* firms. Namely, changes in the set of customers served by a given firm is also a source of volatility in individual sales. We provide evidence that both margins participate to the volatility but that the intensive margin is the most important driver of fluctuations. Namely, only 10% of the *level of* volatilities, whether in the small or in the large, is explained by extensive adjustments. The remaining 80% are attributable to fluctuations at the intensive margin, i.e. the variance of sales for existing seller-buyer pairs. It has to be noted however that, at the individual level, the extensive margin explains about half of the *dispersion* in volatility.

portfolio. With a fat-tailed distribution of producers, this should contribute to reducing the aggregate effect of seller-buyer shocks. In our data however, some large exporting firms happen to be extremely poorly diversified across customers, explaining the strength of these shocks.

As is apparent from our above presentation, this paper contributes to several strands of the literature. The analysis of the volatility in the small brings new evidence to the literature trying to understand the heterogeneity of firms in terms of their volatility.[5] This literature has notably shown that large and/or more experienced firms tend to be less volatile. Indeed, our results confirm this finding. However, we also prove that part of the correlation is due to these firms with a more diversified portfolio of customers that offers natural hedging against buyer-specific shocks.[6] The analysis of the volatility in the large is more closely related to the granularity literature. In this respect, our paper shares a lot with di Giovanni et al. (2014) who use a similar aggregation procedure to analyze the sources of volatility in French data. However, our data structure allows us to pinpoint the role of buyer-specific and match-specific shocks.

The use of export data naturally draws a link with the trade literature. Several contemporary papers also use firm-to-firm trade data to go deeper into the microeconomic structure of aggregate export flows. In particular, Bernard et al. (2014), Eaton et al. (2013) and Carballo et al. (2013) also provide evidence on the heterogeneity of exporters in terms of the number of buyers they interact with. This dimension of heterogeneity is however interpreted in completely different contexts, to discuss the welfare gains from trade (Carballo et al., 2013), individual and aggregate trade patterns (Bernard et al., 2014), or their dynamics (Eaton et al., 2013). Last, a recent and complementary strand of the literature has studied the structure of economic networks as the outcome of an endogenous process (Oberfield, 2011; Chaney, 2014). In contrast to these papers, we take the structure of the network as given and study how it shapes the amount of volatility in trade data.

The rest of the paper is organized as follows. We start with a description of our data and new stylized facts on trade networks in Section 2. In Section 3, we describe our identification strategy of the growth decomposition at the most disaggregated (seller-buyer) level. Next, we present the results in two distinct steps. We discuss the origin of fluctuations *in the small*, at the level of individual firms, in Section 4. Section 5 instead analyzes the question *in the large*, based on aggregate trade flows. Finally, Section 6 concludes.

[5] See, among many others, Comin and Philippon (2006), Thesmar and Thoenig (2011), Decker et al. (2014), or Fort et al. (2013). These papers rely on very different arguments for explaining the heterogeneity in firms' volatility, including the degree of financial constraints (Comin and Philippon, 2006), the heterogeneity in risk-taking behaviors in a context of increasing stock market participation (Thesmar and Thoenig, 2011), the correlation of individual volatility with the age of the firm (Davis et al., 2009), etc.

[6] Note that the argument shares some flavor with Caselli et al. (2015) who also rely on a "diversification" argument for explaining the amount of fluctuations. In their paper however, the argument is applied to the volatility in the large and the way a country's diversification across export markets is offering natural hedging against country-specific shocks.

2 Data and Stylized facts

2.1 Data

The empirical analysis is conducted using detailed export data covering the universe of French firms. The data are provided by French Customs.[7] The full data set covers all transactions that involve a French exporter and an importing firm located in the European Union. Our analysis however focuses on exports to the fifteen "old" members of the European Union, less Greece, Luxembourg, and Austria.[8] For all these countries, we use data for the 1995-2007 period.

Many researchers before us have used individual trade data provided by the French Customs. Most of the time, the data used in empirical approaches are annual data disaggregated at the level of the exporting firm, as in Eaton et al. (2011), Mayer et al. (2014) or Berman et al. (2012), among many others. Some papers also use data detailed at the level of the importer, for instance Blaum et al. (2013). Bricongne et al. (2012) use more disaggregated data, detailed at the level of a transaction: For each exporting firm, they know how many times a good has been exported to one destination in the year under consideration. In comparison with this paper, our data are even richer since we know, among other characteristics, the identity of the exporting firm *and* the identity of the importer it serves.

For each transaction, the data set specifies the identity of the exporting firm (its name and its SIREN identifier), the identification number of the importer (an anonymized version of its VAT code), the date of the transaction (month and year), the product category (at the 8-digit level of the combined nomenclature) and the value of the shipment. In the analysis, data will be aggregated across transactions within a year, for each exporter-importer pair. This helps focus on the most important novelty in the data, which is the explicit identification of both sides of the markets, the exporter and its foreign partner.[9]

While goods are perfectly free to move across countries within the European Union, firms selling goods outside France are still compelled to fill a Customs form. These forms are used to repay VAT for transactions on intermediate consumptions. This explains that the data are exhaustive. One caveat, though: small exporters are allowed to fill a "simplified" form that does not require the product category of exported goods. This is problematic whenever the empirical strategy controls for sector-specific determinants of the outcome variable

[7]In particular, we would like to thank Thierry Castagne who took time to explain the specificities of the data.

[8]The reason for leaving these three countries aside comes from the difficulty, not to say the impossibility, of identifying individual buyers for these destinations. We found breaks in the panel dimension of buyers' identity. We also had to exclude from the analysis the new member states of the European Union because the time dimension was too short in their case.

[9]Notice that, even though we track each sale a seller makes to each country, we cannot do the same for buyers. More precisely, we cannot know if the same buyer buys from two foreign sellers from two different countries. More generally, since we do not have additional information on the buyer, we cannot say whether it is an affiliate of the same (multi-national) firm as the seller or indeed if two buyers in our data are connected through multinational linkages.

since the corresponding transactions cannot be included in the data set. The "simplified" regime concerns firms with total exports in the European Union in a given year below 100,000 euros (150,000 euros since 2006). Put differently, some of our regressions do not include the smallest exporting firms. Since the analysis focuses on "granular" fluctuations, which mostly involve large firms in an economy, we believe this absence to be immaterial. We however checked that the most important stylized facts still prevail if we include the small firms, without controlling for the product dimension.

Given the quality of the data, little cleaning is necessary to construct the final data set. There is only one type of flows that we remove. In some cases, the country code is not consistent with the country code that can be recovered from the importer's identifier. This happens when a French firm plays the role of an intermediary to sell a good produced in a given country bought by a customer in another country. Since such transactions cannot be qualified as "French exports" stricto sensu, they are removed from the database.

In 2007, we have information on 42,888 French firms exporting to 334,905 individual buyers located in the 11 countries of the European Union. Total exports by these firms amounts to 207 billions euros. This represents 58% of French total exports. Detailed summary statistics by destination country are provided in Table 1. While large destination markets naturally involve more firms on both sides of the border, the density of trade networks, as measured by the number of active pairs divided by the potential number of relationships, is instead lower in countries like Germany or Belgium.

– Table 1 about here –

The firm-to-firm data are used to describe the structure of trade networks, in the cross-section. We also use the time-dimension to compute measures of sales growth at different levels of aggregation. Let us denote the growth rate as g_{sbt}, g_{st}, g_{bt} and g_t, respectively for the seller-buyer, the seller, the buyer and the aggregate levels. In everything that follows, the analysis is conducted destination by destination but we simplify the notations by omitting the index for the destination market.[10] g_{st} and g_t are thus defined as the growth of sales to a specific destination, measured for one specific seller or in the aggregate. Our measure of volatility is the variance of annual growth rates computed at each level. We restrict our attention to the subset of second-order moments computed on at least four points.[11] Finally, to minimize the effect of outliers on

[10]This choice is motivated by our willingness to isolate the role of diversification across buyers within a destination from that of diversification across destination markets. The role of diversification across destinations is pointed out in Caselli et al. (2015).

[11]Whereas four points may seem a small number for computing measures of volatility, this restriction reveals itself rather constraining once one realizes that it relies on the number of years a given firm serves a specific destination or even the number of years a given seller-buyer pair is active. Figure A.1 displays the frequency of pair durations in our data source, where the duration of a pair is defined as the number of growth rates the sample contains for that seller-buyer pair. The low average duration of relationships also explains why we do not compute time-varying measures of volatility based on sub-periods, as is often done in the macroeconomic literature.

our measures of volatility, we base our estimates in Section 3 on observations for which the seller growth rates lies in the interval $[-0.8; 4]$. Table A.1 in Appendix quantifies how restrictive these constraints are, as measured by the sample coverage.

In addition to firm-to-firm trade flows, we merge historic information on firms' participation in foreign markets using the firm's Siren identifier. The data, provided by the French Customs, go back to 1990. This allows us to to construct a measure of firm's "experience" in each destination market. In Chaney (2014), it is the history of the firm in a destination that explains how many contacts it has there. Even though we do not seek to explain the structure of firms' trade networks, trade experience will be used as a control variable in some of our regressions. To better summarize experience, we will contrast "entrants" (experience below two years), "young exporters" (experience between 2 and 5 years), and "mature exporters" (experience of at least 5 years).

Furthermore, we complement the above data sets with the so-called LiFi. The data, built by the French Statistical office (INSEE), provide information on ownership relations between French firms and their affiliates, in France as well as abroad. We merge the LiFi using again the Siren identifier and measure whether exporters in our data are entities of a foreign multinational or if they own foreign affiliates in the market where they export goods. This allows to control for the impact of intra-firm linkages on the structure of trade networks and the magnitude of firm-level fluctuations.[12] Even though the LiFi is not exhaustive, it has a satisfactory coverage, as discussed in Kleinert et al. (2012).

2.2 Stylized facts on trade networks

In this section, we describe the structure of French firms' trade networks, as of 2007. We first describe the distribution of trade flows *across* firms. We coin it the diversification in the large. We then present the diversification in the small. This corresponds to the distribution of trade flows within firms across trade partners. Finally, we discuss the importance of the extensive margin for the growth of individual and aggregate exports.

Diversification in the large. The skewness of individual sales is key to generate granular fluctuations (Gabaix, 2011). If the distribution of sales were symmetric, idiosyncratic volatility would have a negligible impact on aggregate fluctuations since individual shocks would compensate each other. In Gabaix (2011), it is thus the distribution of sales across firms that determines the prevalence of idiosyncratic supply shocks in aggregate fluctuations. As argued in the introduction, a similar reasoning applies if idiosyncratic shocks hit the buyer side of the economy. Then what matters is the distribution of purchases across

[12]The French Customs do not distinguish between intra-firm and arm's length trade flows. Information on the nationality of the firm's affiliates and/or headquarter is used as indicative of the possibility that part of the bilateral trade flows observed in our data are indeed intra-firm.

buyers. Finally, if shocks are specific to a seller-buyer transaction, the distribution of sales across transactions is key to understand the aggregate impact of the shocks.

The concentration of exports across sellers is illustrated in Figure 1, first panel. The figure displays the cumulative value of sales across firms of increasing size.[13] It confirms a well-known stylized fact of the trade literature, namely that the distribution of sales across exporting firms is extremely skewed. At the top of the distribution, 1% of firms are responsible for about 55% of exports.

– Figure 1 about here –

This extreme skewness shows up in the Herfindahl index, equal to .005 in our data (See Table 2, column (2)).[14] While the absolute number is not meaningful, it represents 234 times the Herfindahl one would observe in a counterfactual world with S exporters symmetric in size ($Herf/(1/S) = 234$). This ratio varies significantly depending on the destination country under consideration (column (3), Table 2). It is maximum for French sales in Spain and minimum for exports to Germany.

– Table 2 about here –

Imports from France are also extremely concentrated across buyers, as illustrated in Figure 1, second panel. At the top of the distribution, 1% of firms are responsible of about 70% of imports. The Herfindahl index of import purchases is equal to .001 and is almost 500 times larger than it would be under a uniform distribution (Table 2, columns (6) and (7)). This implies that shocks to the largest importers in each destination are very likely to show up in aggregate fluctuations.

The skewness of either export or import distributions has already been documented in the trade literature. There is however no such evidence regarding the distribution of sales across exporter-importer pairs. This is what the third panel in Figure 1 and the last four columns in Table 2 illustrate. Here as well, the distribution is highly skewed: the 1% largest transactions represent 70% of aggregate trade and the Herfindahl index is equal to .001. In a (large) destination country such as Spain, the ten largest transactions account for 22% of French exports. This proportion is even larger in smaller destinations such as Ireland and Sweden. The least concentrated country is Germany. In this country among the 249,196 firm-to-firm pairs exchanging in 2007, the ten largest account for "only" 9% of total French exports.

Whatever the dimension considered, the data show an extreme degree of concentration. Aggregate trade is not diversified in the large and is thus likely

[13] Here, the size of a firm is measured by the total value of its sales in the 11 EU countries that we consider in the paper. Results are unchanged if we instead focus on the distribution of destination-specific sales.

[14] As expected, trade is more granular than total sales. Indeed, di Giovanni et al. (2014) reports a Herfindahl of sales for French manufacturing firms of .0035. The distribution of sales across French exporters is almost twice as concentrated as the distribution of total sales.

to display granularity. Next, we show that trade is also poorly diversified in disaggregated data.

Diversification in the small. In presence of firm-specific shocks, the magnitude of fluctuations in individual sales depends in particular on the structure of individual exporters' clientele.[15] We now describe the extent of sellers' diversification - measured by the number of customers within each French exporter's portfolio.

– Figure 2 about here –

Figure 2 presents the distribution of the number of buyers in each French exporter's portfolio. We focus on the number of buyers within a given destination country. The top panel of Figure 2 (circles line) represents the share of sellers having at least a given number of buyers in a given destination. The bottom panel (circles line) presents the share of these firms in total exports, i.e. taking into account the heterogeneity across sellers of different size. In 2007, 43% of French sellers export to a single buyer (top panel). These sellers are exposed to a maximum level of idiosyncratic demand risk since they are not diversified at all across buyers within a destination. Such sellers only account for 18% of total sales, however (bottom panel). At the other side of the distribution, 12% of firms have more than 10 partners in their typical European market. These firms are large on average since they represent around 40% of total exports. These distributions thus reveal a large amount of heterogeneity in the level of diversification of French exporters, with larger exporters selling to more buyers.

The number of clients in a firm's portfolio is not sufficient to fully assess a seller's diversification. Indeed, a firm may have many partners but be extremely poorly diversified if most of these partners buy tiny amounts. This possibility seems consistent with our data, as shown by the additional lines displayed in Figure 2. While the green circles line shows the number of buyers using the total sales of each exporter, the other lines restrict the analysis to a given amount within each firm's exports. Namely, for each exporting firm, buyers are ranked according to their (decreasing) size and the number of clients is computed by excluding from the computation the smallest buyers representing some given share of exports. Using this strategy, we see that, among the 12% of firms that serve more than 10 buyers, many serve tiny importers with cumulative share less than 10% of the firm's exports. Once such tiny buyers are removed, only 6% of sellers are found to serve at least 10 partners. This number is close to 0 when one concentrates on only half of the firm's sales. These findings indicate that exporters' sales are not well diversified across buyers: even large firms with a rich portfolio of clients tend to concentrate their sales on one or two "main" partners.

– Table 3 about here –

[15] See Section 4 for a detailed description of the mechanisms surrounding this relationship.

We complete our characterization of individual firms' degree of diversification using a multivariate regression framework. Results are presented in Table 3. We explain the variance across firms in their degree of diversification using a set of observable characteristics. First, we include indicators to control for all determinants of diversification that are specific to a sector in a destination (columns (1), (3), (4) and (6)) and/or specific to the exporter (columns (2), (3), (5) and (6)). Columns (1)-(3) use the number of clients in the exporter's portfolio as left hand side variable and columns (4)-(6) use the Herfindahl index. Since both variables are negatively correlated, we expect the estimated coefficients to be of opposite sign.

The first explanatory variable introduced in the regressions is the size of the exporting firm, as measured by its export sales in the destination. We also include its square to capture non-linear effects. The coefficients obtained with these two variables imply a non-linear relationship between the firm's size and the diversification of its portfolio: Large firms tend to be better diversified, up to a threshold above which the diversification of individual firms decreases.[16] Following Chaney (2014), we also control for the experience of the firm in the destination.[17] Consistent with the expectation, we find that the number of buyers served by a given exporter and the diversification of its sales (as measured by the inverse of the Herfindahl) are increasing in its experience.

The second set of explanatory variables captures the firms diversification across products rather than across buyers. Columns (1)-(3) regress the number of clients on the number of 8-digit products the firm has in its portfolio while the Herfindahl index of sales across products is used in columns (4)-(6). If diversification across buyers was entirely due to the firm expanding its range of products, the coefficient on these variables would be equal to one, which is not what we find. Even though the estimated coefficient is very significant and positive, it is far away from one implying that firms diversify across buyers, even within products.

Our regression exercise also controls for multinational linkages that the French exporter might share with firms in the export destination. This will happen if the firm is an affiliate of a foreign MNE located there (variable called "$\mathbb{1} = 1$ if HQ from dest.") or has affiliates in the destination (variable called "$\mathbb{1} = 1$ if affiliates in dest."). Estimated coefficients suggest that export flows are less diversified across buyers when the exporter is involved in a multinational relationship with firms in the destination. This might be expected if part of the recorded exports are in fact intra-firm trade, a form of trade that does not expose partners to the same type of risks.

The last element of this table has to do with the heterogeneity across exporters in their *potential* for diversification. Depending on the type of products it sells, a firm may indeed face a very large number of potential clients or an

[16]Increasing size has a positive impact on firm's diversification for exports below the 80^{th} percentile. Then, increasing size is associated with a reduction in the level of diversification.

[17]In Chaney (2014), exporters gradually accumulate customers in each destination through random meetings. As a consequence, the number of buyers in a firm's portfolio increases over time.

oligopsonic demand. Such differences in their *potential* for diversification may well be at the root of the heterogeneity in the *actual* degree of sales diversification. We thus control for the number of potential clients that an exporter faces.[18] As expected, the variable is positively correlated with the number of buyers in the firm's portfolio, but the correlation is far from perfect. The correlation is also positive, but small, between the *actual* and the *potential* Herfindahl index of sales in column (4)-(6).[19]

Together, these results suggest that the degree of sales diversification is strongly heterogeneous across firms and systematically correlated with characteristics of the firm, namely its size, the number of potential clients it faces and its experience as an exporter. The coefficients on multinational linkages and the one on firm's size suggest that the largest firms are not the most diversified ones. Individual shocks to transactions in which such firms are involved are thus likely to have sizable aggregate implications.

To complete our analysis of diversification in trade networks, in the small, we look now at the way individual importers spread their purchases across French exporters. Figure 3 presents the distribution of the number of French sellers within each importer's portfolio for the year 2007. Here as well, the distribution is expressed in relative terms with respect to the population of importers (top panel) and to the total value of trade (bottom panel).

– Figure 3 about here –

In 2007, 65% of foreign buyers trade with a single French exporter (top panel).[20] These importers are however small, on average: They represent only 16% of total sales (bottom panel). At the other side of the distribution, 4% of firms have more than 10 French partners. These firms are large on average since they represent around 40% of total imports. Interestingly enough, the basket of foreign buyers is extremely concentrated, even for buyers interacting with several French exporters. Almost 90% of importers buy 90% of their imports from France from a single firm. They account for almost 70% of exports.

The extensive and intensive margins of exports. The analysis so far has been static. The patterns of the distribution of sales and purchases across and within firms exhibit little change over the 1995-2007 period. This is true in the aggregate but, in disaggregated data, a substantial share of the action takes place at the extensive margin. At the seller-level, individual growth rates

[18]The variable is calculated based on the observed number of buyers that purchase one specific type of products. Using the firm's observed portfolio of products and the number of potential buyers for each of those products, it is possible to compute the theoretical number of buyers that an exporter could serve, i.e. the supply of clients it is offered.

[19]Here, the potential Herfindahl is calculated as before and assuming that the firm allocates its sales across all potential buyers in proportion to their relative demand.

[20]A small proportion of these firms interact with sellers which themselves are not diversified in sales. Namely, 1.5% of all transactions in the 2007 data involve an exporter serving a single buyer which itself does not interact with any other exporter. Such "1-to-1" trade flows are mechanically excluded from the identification strategy in Section 3.

are in part driven by the net entry of buyers in firms' portfolio of customers. At the aggregate level, the effect is reinforced by entries and exits of sellers into different destination markets. To assess the economic importance of such adjustments, we compute the relative contributions of the intensive and the extensive margins to the annual growth of sales. Details on the decomposition are provided in the Appendix of the paper, Section A.

Consider first the growth rate of individual sales, by destination. At this level of aggregation, extensive adjustments are attributable to the net entry of customers in the firm's portfolio. Following the same logic as in Feenstra (1994), we have:

$$g_{st}^{Tot} = g_{st} + g_{st}^{Ext.} \tag{1}$$

where g_{st}^{Tot} is the total growth rate of exports of firm s in a given destination. g_{st} denotes the intensive component of this growth rate, computed on the subset of incumbent buyers (the set $B_s \equiv B_{st} \cap B_{st-1}$ of buyers with a strictly positive demand to s in both $t-1$ and t). Finally, $g_{st}^{Ext.}$ is the extensive component, attributable to the net entry of buyers into the firm's portfolio.[21]

The median and the mean contributions of these two terms to the growth of individual destination-specific sales are reported in the first and second columns of Table 4. The lion's share of individual growth is clearly driven by the intensive margin. The mean contribution of the intensive margin is 80% and its median contribution 100%.

– Table 4 about here –

Columns (3) and (4) of Table 4 present the contribution of the extensive and intensive margins to the growth of *aggregate* exports. Compared to equation (1), a third term appears, which accounts for the net entry of sellers in a market (i.e. variations in the set S_t of exporters serving the destination in a given period). The decomposition now writes as:

$$g_t^{Tot} = g_t + g_t^{Ext-buyer} + g_t^{Ext-seller} \tag{2}$$

in which g_t^{Tot} represents the overall growth of aggregate exports to a destination, which decomposes into i) an intensive component g_t computed on "intensive" transactions (the set of seller-buyer pairs such that $(s,b) \in \bigcup_{s \in S} B_s$ where $S = S_t \cap S_{t-1}$ is the set of "intensive" exporters), ii) a buyer component of the extensive margin $g_t^{Ext-buyer}$ due to the net entry of new buyers into incumbent sellers' portfolio, and iii) a seller component of the extensive margin $g_t^{Ext-seller}$ due to the net entry of sellers into the market.

Again, the intensive margin explains most of the aggregate export growth. Its contribution is nonetheless smaller than at the individual level. The mean and median contributions of the intensive margin are both equal to 0.65. The relative size of the two extensive components is sensitive to the way we measure

[21] By definition, we cannot compute the decomposition in absence of intensive margin (g_{st} and $g_{st}^{Ext.}$ not defined). This reduces slightly the number of observations used in the analysis as shown in Table A.1.

their contribution. Looking at the median, we find that the net entry of sellers contributes to 15% of aggregate growth while the net entry of new buyers accounts for 12%. Looking at the mean, the contribution of the net entry of sellers is negative (-5%) while the contribution of new buyers reaches 40%.[22]

These results make it clear that the lion's share of export growth is driven by the intensive margin – the growth of existing seller-buyer relationships. In the rest of the analysis, we will study the volatility of trade neglecting extensive adjustments. Results in Appendix A however complete the analysis with a quantification of the contribution of extensive adjustments to the overall volatility of sales, both in the small and in the large.

3 Empirical strategy

Having discussed the main statistical properties of the trade networks under study, we now turn to the actual analysis of the dynamics of trade in those networks. As explained in the introduction, we start from the most disaggregated level and then aggregate up, step by step. In this section, we discuss our strategy for identifying the sources of shocks at the most disaggregated level. This strategy is motivated by a theoretical framework, which we use to identify the orthogonality condition later exploited.

3.1 Sources of firm-to-firm trade growth

In this section, we develop a partial equilibrium model of the demand for imported goods, in which we put a variety of fundamental shocks to obtain rich predictions regarding the determinants of disaggregated trade growth.

The demand side of the model is represented for a buyer b which produces a consumption good using various imported inputs and sells it to a representative consumer. The technology for producing y_b units writes as follows:

$$y_b = \left[z_b \sum_s \left(z_{sb} x_{sb} \right)^{\frac{\sigma-1}{\sigma}} \right]^{\frac{\sigma}{\sigma-1}} \tag{3}$$

where x_{sb} is the demand for the input produced by a seller s, $\sigma > 1$ is the elasticity of substitution between input varieties, z_{sb} is a preference parameter for input s and z_b is a measure of the buyer's productivity. Since we will ultimately apply the model to trade data between French exporters and various European buyers, it is assumed that buyer b uses as sole inputs intermediate goods supplied by French firms. While unnecessary, it is of course straightforward to adjust the model to a more general technology function combining other inputs as well as labor and capital.

[22]It is worth noting that aggregate contributions are computed over 11 countries and 11 years. The results for the mean are driven by a few outliers. Excluding these outliers gives numbers which are comparable to the median results.

Given this production function, the program of the buyer consists in the minimization of total costs induced by the production necessary to satisfy the demand of the market:

$$y_b = p_b^{-\eta} A$$

where p_b is the price charged by the buyer to her representative consumer and A an aggregate demand shifter (which will later be allowed to be sector-specific). $\eta > 1$ measures the price elasticity of the final demand. The CES demand function implies that the buyer charges her consumer with a price p_b which is a constant mark-up $\frac{\eta}{\eta-1}$ over her marginal cost c_b.

Solving the cost minimization program implies the following (nominal) demand addressed to supplier s:

$$p_{sb}x_{sb} = \left(\frac{p_{sb}}{z_{sb}c_b}\right)^{1-\sigma} c_b \left(\frac{\eta}{\eta-1}c_b\right)^{-\eta} A z_b^{\sigma} \tag{4}$$

where p_{sb} denotes the price charged by supplier s on her customer b. In equilibrium, the marginal cost writes:

$$c_b = z_b^{\frac{-\sigma}{\sigma-1}} \left[\sum_s \left(\frac{p_{sb}}{z_{sb}}\right)^{1-\sigma}\right]^{\frac{1}{1-\sigma}}$$

To obtain a demand equation expressed in terms of the fundamentals, we finally need to define the pricing strategy of input providers. It is assumed that varieties of inputs are produced using a production function which is linear in labor. The productivity of labor is assumed to be a function of an aggregate productivity component Z and a firm-specific component z_s. Taking the cost of labor in France as the numéraire and assuming that input providers compete under monopolistic competition, the price set by exporter s for her sales to buyer b is:[23]

$$p_{sb} = \frac{\sigma}{\sigma-1}\frac{1}{z_s Z} \tag{5}$$

Plugging equation (5) into (4) and taking the first difference in logs over time gives the predicted equation for the dynamics of trade in the model:

$$\begin{aligned} g_{sbt} &\equiv d\ln p_{sbt}x_{sbt} \\ &= (\sigma-1)d\ln Z_t + d\ln A_t + (\sigma-1)d\ln z_{st} \\ &\quad +(\sigma-\eta)d\ln c_{bt} + \sigma d\ln z_{bt} + (\sigma-1)d\ln z_{sbt} \end{aligned} \tag{6}$$

[23] Here, it is assumed that there is no delivery cost charged on international sales to buyer b. Since we later focus on the dynamics of trade, it would be exactly equivalent to assume that there is such delivery cost but it is constant over time. In the empirical model, variations in the transportation cost would be absorbed into the aggregate shock (if they are common across firms within a sector), or the match-specific shock (if they are specific to the firm and/or her customer).

where the growth of the marginal cost can be written using a Taylor approximation:

$$\begin{aligned} d\ln c_{bt} &= -\frac{\sigma}{\sigma-1} d\ln z_{bt} + \sum_s w^b_{st-1} d\ln\left(\frac{p_{sbt}}{z_{sbt}}\right) \\ &= -\frac{\sigma}{\sigma-1} d\ln z_{bt} - d\ln Z_t - \sum_s w^b_{st-1}(d\ln z_{st} + d\ln z_{sbt}) \end{aligned}$$

with w^b_{st-1} the share of seller s in buyer b's input cost, in $t-1$.

Equation (6) thus defines the dynamics of trade as a function of the growth of the fundamentals, namely the two supply parameters, z_{st} and Z_t, and the three demand parameters, z_{sbt}, z_{bt} and A_t.[24] The last step consists in specifying how these parameters evolve over time. In what follows, it will be assumed that their dynamics is driven by i.i.d. shocks. Namely:

$$\begin{aligned} Z_t &= Z e^{\varepsilon_{Zt}}, \quad \varepsilon_{Zt} \sim \mathcal{N}(0, \sigma^2_Z/2) \\ A_t &= A e^{\varepsilon_{At}}, \quad \varepsilon_{At} \sim \mathcal{N}(0, \sigma^2_A/2) \\ z_{st} &= z_s e^{\varepsilon_{z_s t}}, \quad \varepsilon_{z_s t} \sim \mathcal{N}(0, \sigma^2_{z_s}/2) \\ z_{bt} &= z_b e^{\varepsilon_{z_b t}}, \quad \varepsilon_{z_b t} \sim \mathcal{N}(0, \sigma^2_{z_b}/2) \\ z_{sbt} &= z_{sb} e^{\varepsilon_{z_{sb} t}}, \quad \varepsilon_{z_{sb} t} \sim \mathcal{N}(0, \sigma^2_{z_{sb}}/2) \end{aligned}$$

Plugging this into equation (6) delivers the structural equation of the dynamics of trade which will later be estimated, namely:

$$\begin{aligned} g_{sbt} &= (\eta-1)d\varepsilon_{Zt} + d\varepsilon_{At} + (\sigma-1)d\varepsilon_{z_s t} + \frac{\sigma}{\sigma-1}(\eta-1)d\varepsilon_{z_b t} \\ &\quad +(\sigma-1)d\varepsilon_{z_{sb} t} - (\sigma-\eta)\sum_s w^b_{st-1}(d\varepsilon_{z_s t} + d\varepsilon_{z_{sb} t}) \end{aligned} \tag{7}$$

We will now explain how we exploit the structure of the data to estimate the different components of the above equation.

3.2 Identification strategy

The easiest way to visualize our data is, at each date t, as a bipartite graph with sellers and buyers as the two types of nodes.[25] An edge in this graph is

[24] The classification of shocks into the supply and demand categories is somewhat arbitrary in this context. In what follows, all shocks affecting the buyers are called demand shocks, even though they can be driven by changes in the productivity of those firms. From the point of view of the French exporter, such shocks induce an exogenous change in the demand for exports, which justifies the choice of this vocabulary.

[25] This structure would also have another dimension if data were not aggregated across products. In that case, the nodes at each side of the graph would be sellers and buyers of a specific product.

therefore a (positive) sale between two nodes, one of each type. Equation (7) describes how the magnitude of such edge evolves over time, in our model.[26]

In the following, we use the richness of the network structure to distinguish between the different components of growth introduced into the model. We take inspiration from the labor literature on the dispersion of wages, notably Abowd et al. (1999).[27] Following Abowd et al. (1999), the determinants of the dispersion of wages in employer-employee linked data are estimated using high-dimensional fixed effect estimators. Our model also delivers such fixed effect decomposition. After some rewriting, Equation (7) indeed rewrites:

$$g_{sbt} = f_{Ct} + f_{st} + f_{bt} \underbrace{- \frac{\sigma - \eta}{\sigma - 1} \sum_s w^b_{st-1}(f_{st} + \nu_{sbt})}_{BSIC_{bt}} + \nu_{sbt} \tag{8}$$

where:

$$\begin{aligned} f_{Ct} &= (\eta - 1)d\varepsilon_{Zt} + d\varepsilon_{At} \\ f_{st} &= (\sigma - 1)d\varepsilon_{z_s t} \\ f_{bt} &= \frac{\sigma}{\sigma - 1}(\eta - 1)d\varepsilon_{z_b t} \\ \nu_{sbt} &= (\sigma - 1)d\varepsilon_{z_{sb} t} \end{aligned}$$

are the fixed components and the match-specific residual explaining the cross-sectional dispersion of growth rates.

The partial equilibrium model of Section 3.1 thus delivers a decomposition of the firm-to-firm trade growth into three terms, a macroeconomic component, a seller-specific term and a buyer-specific effect, plus a residual term which is specific to the seller-buyer match. The buyer-specific part is itself composed of two terms, the buyer fixed effect f_{bt} and a weighted average of the seller- and match-specific shocks affecting her partners ($BSIC_{bt}$). The $BSIC_{bt}$ term arises from the buyer-specific input cost index being responsive to the price adjustments hitting each of her partners. A negative productivity shock to seller s thus increases the input cost which smooths the direct impact that the shock has on the demand addressed to seller s. Moreover, this generates externalities

[26] Another dimension that the model does not take into account is the extensive margin. Fundamental shocks might indeed affect the existence of an edge: the two nodes exist but the buyer does not buy from this seller at date t whereas it did at date $t-1$ or instead starts buying at date t while it was not at date $t-1$. The shock may also affect the existence of a node, e.g. a sufficiently bad productivity shock might force some sellers out of the market. Neither our model nor the estimation strategy explicitly take extensive adjustments into account. Appendix A discusses into details the extent to which this might bias our results. We argue that, because the extensive margin contributes little to the overall variance of sales, the potential bias is small.

[27] This literature seeks to decompose the total cross-sectional distribution of wages into the shares which are attributable to firm-specific variables, worker-specific elements and match-specific determinants. To this aim, high-dimensional fixed effect estimators are applied to employer-employee linked data. Identification is achieved thanks to the connectivity in employer-employee networks induced by workers' mobility across firms.

on the rest of the buyer's portfolio of partners. The presence of this buyer-specific input cost component in equation (8) is important because it generates a negative correlation between the buyer-specific term and the match-specific residual. Absent the correlation, equation (8) could be estimated using Abowd et al. (1999).

Their strategy is inconsistent, however, because of the correlation between the input cost index and the residual in equation (8). In order to solve the problem, the Abowd et al. (1999) estimator is instead applied to a modified version of equation (8) in which the buyer-specific input cost index no longer appears:

$$\tilde{g}_{sbt} = (1+\lambda) f_{Ct} + f_{st} + (1+\lambda) f_{bt} + \nu_{sbt} \tag{9}$$

where:

$$\begin{aligned}
\tilde{g}_{sbt} &= g_{sbt} + \lambda g_{bt} \\
g_{bt} &= \sum_s w^b_{st-1} g_{sbt} \\
\lambda &= \frac{\sigma - \eta}{\eta - 1}
\end{aligned}$$

In matrix format:

$$G_t = \alpha_t f^C_t + \chi_t f^S_t + \beta_t f^B_t + \nu_t$$

where G_t is the vector that contains the $\tilde{g}_{sbt}$ terms ($N_t \times 1$, where N_t is the number of observations for year t), α_t is the design matrix for the year-t country-sector effects ($N_t \times N^C_t$, where N^C_t is the number of country-sector for year t),[28] χ_t is the design matrix for the year-t seller effects ($N_t \times N^S_t$, where N^S_t is the number of sellers for year t), β_t is the design matrix for the year-t buyer effects ($N_t \times N^B_t$, where N^B_t is the number of buyers b, at date t), and ν_t is the vector of residuals ($N_t \times 1$).

Given a value for λ, the components of equation (9) can be identified in the cross-section of year-specific growth rates. Identification is achieved assuming:

$$E(\nu_{sbt} | (s,t); (b,t)) = 0 \tag{10}$$

or, in matrix format:

$$E(\nu_t | \chi_t, \beta_t) = 0 \tag{11}$$

These assumptions are identical to those in Abowd et al. (1999). Notice here that there is no explanatory variable, X, in the above model (except the α_t, the year-industry effects). To reach identification of the seller and buyer components, the buyers and sellers must be connected in the sense of belonging

[28] Because some of our firms sell multiple products, the definition of the firm's "industry" is not necessarily straightforward. We chose to affect each seller-buyer pair to the "industry" that corresponds to the most important product constituting the corresponding trade flow. Industries are defined the the 2-digit level of the HS nomenclature.

to a connected group (Abowd et al., 2002). In particular, all pairs of buyers and sellers for which both the buyer and the seller have a unique partner cannot have identified effects, hence unconnected pairs of partners. They represent about 4% of the observations in the regression sample. For each connected group, all the buyer and seller effects but one are identified.[29]

Applying the above strategy requires that we take a stand on the value of λ, a function of the two demand elasticities of the partial equilibrium model, namely σ the elasticity of substitution between French inputs, and η the price elasticity of demand addressed to foreign buyers. Up to now, we have not put any restriction on these parameters except that they are both larger than one, and homogeneous across buyers. We recover this function of the parameters using an additional exogeneity condition suggested by the theoretical model, namely the orthogonality between the seller- and the buyer-specific components:

$$E(f_{st} f_{bt}) = 0 \tag{12}$$

Under the "true" value for λ, condition (12) should apply. We thus implement a grid-search algorithm on all the possible values for λ and pick the value which best satisfies the model-implied orthogonality condition. Note that the algorithm is straightforward to implement because there is a monotonous relationship between the value chosen for λ and the magnitude of the correlation of the seller and buyer fixed effects. Appendix B gives more details on the estimation procedure.[30]

To summarize, the estimation procedure consists in:

1. Estimating the components of equation (9) under the exogeneity condition (10) for different values of λ

2. Applying a grid-search algorithm to choose the value for λ which best matches the orthogonality condition (12), and the corresponding estimated components,

3. Recovering the theoretical decomposition in equation (8) which measures the relative contribution of different types of shocks in driving the cross-sectional distribution of firm-to-firm growth rates.

[29]In a previous version of this paper, we adopted a slightly different version of the model, in which seller effects where country-specific. Therefore, the equation above could be estimated country by country. In this version, a seller is endowed with a unique seller-effect in the year common to all its destinations. Obviously, buyer are all country-specific since there is no common identifier. The benefits of this new strategy are clear on at least two grounds. First, the network is denser and many more observations are now "connected"; hence with identified effects. Second, because we have more available observations to estimate each of the seller effects (at least for those sellers that export to at least two countries), the precision of the estimated seller effects is increased (Abowd et al., 2002).

[30]As detailed in Appendix B, the orthogonality condition (12) relies on the asymptotic properties of the model. Given that the actual network is relatively sparse, it might be that we do not achieve orthogonality between the seller and buyer components because of measurement errors on some of the components of the network. We take into account this potential "limited mobility bias" in the estimation of λ. Namely, we quantify the magnitude of the bias using a numerical simulation and target a value for the correlation which is consistent with the results of the simulation. See details in Appendix B.

Results for the decomposition are presented in Tables 5 and 6. They are obtained for $\hat{\lambda} = .76$ which is consistent with the price elasticity of demand for French inputs being slightly above the price elasticity that buyers face on their own market (eg. 0.76 is consistent with $\eta = 3$ and $\sigma = 4.5$). Such value is consistent with the view that markups increase along the production chain.

Table 5, columns (1)-(3), reports the mean effects, their standard deviations and the number of estimated components. Column (4) then reports the median contribution of each component to the overall growth *level* while column (5) is a partial correlation coefficient which interprets as the contribution of the component to the cross-sectional *dispersion* of firm-to-firm growth rates. Table 6 reports the full correlation table of the various estimated effects. Notice that the correlations are not corrected for the part due to estimation errors.

– Tables 5 and 6 about here –

The number of growth rates for which we can identify all three individual effects is almost equal to 3.2 millions. There are 12 years, 11 countries and 35 2-digit industries, hence almost 4,000 macro shocks. Finally, we are in position to identify more than 200,000 seller (time) effects, using an average of 7 observations per effect and 800,000 buyer (time) effects, using in average slightly less than 4 observations per effect. Without much surprise, the residual match-specific component is the most important component, explaining more than 60% of the level *and* dispersion of firm-to-firm growth rates. The other two individual components, namely the seller-specific and the buyer-specific terms, also contribute substantially to the heterogeneity in the data, respectively accounting for 12 and 25% of the dispersion.

In Table 6, we see that the residual is indeed orthogonal to the buyer and the seller effects, as hypothesized in assumption (10). The correlation of the sellers' and the buyers' effects is around -.07 which is also what was obtained in the "fake" sample in which we simulated orthogonal shocks given the actual network structure of the data. Based on Abowd et al. (2002), and Andrews et al. (2008), we interpret this negative correlation as spurious due to a "low-mobility bias" and use it as input in the grid search procedure to estimate λ.[31] Finally, the negative correlation between f_{st} and $BSIC_{bt}$ is consistent with expectations. A positive productivity shock to seller s drives the input cost index of her partners down which partially counteract the direct effect that the shock has on the demand for this seller's input. Note that this effect also generates spatial correlation within the set of French sellers connected to the same buyer. The negative correlation between f_{bt} and $BSIC_{bt}$ is not consistent with the theoretical model. It is a side-effect of the "limited mobility bias".

[31] An estimation error in the buyer effect translates into an estimation error of the opposite sign in the seller effect. Such estimation errors are more likely to exist in not-very-dense networks. In addition, Andrews et al. (2008) show that the absence of observable controls (X) magnifies this bias.

3.3 Results

To understand the structure of volatility of transactions between buyers and sellers, using the estimated effects presented just above, we compute the volatility of the growth rate of sales for all seller-buyer pairs using all relevant observations.[32] We also compute the volatility of the seller (resp. buyer, resp. seller-buyer, resp. macro) effects using all estimated effects obtained from the decomposition (8) and the associated covariance terms. The variance of seller-buyer growth rates thus decomposes as follows:

$$\begin{aligned} Var(g_{sbt}) &= Var(f_{Ct}) + Var(f_{st}) + Var(f_{bt}) + Var(\nu_{sbt}) \\ &\quad + Var\left(BSIC_{bt}\right) + Cov \end{aligned} \tag{13}$$

where $Var(x_t) \equiv \frac{1}{T}\sum_t (x_t - \bar{x})^2$ is our measure of volatility which is here computed for each seller-buyer pair. Cov is a sum of covariance terms involving the different growth components. Under the assumptions of the model:

$$\begin{aligned} Cov &= Cov\left(f_{st} + \nu_{sbt}, BSIC_{bt}\right) \\ &\approx \frac{\eta - \sigma}{\sigma - 1} w_s^b \left[Var(f_{st}) + Var(\nu_{sbt})\right] \end{aligned}$$

where the approximation in the second line comes from the assumption that the w_{st-1}^b weights are constant over time.

– Table 7 about here –

Summary statistics on the decomposition are presented in Table 7. Columns (1) and (2) respectively report the mean and standard deviations of the variance components, computed over the population of firm-to-firm relations. On average, the sum of covariance terms is slightly negative, which is consistent with the model, in which $BSIC_{bt}$ is negatively correlated with the seller- and match-specific components when $\eta < \sigma$. Column (3) in Table 7 reports the median contribution of each variance component to the overall level of firm-to-firm volatility. Finally, Column (4) reports their contribution to the dispersion of firm-to-firm volatility. Note that, at that level of disaggregation, the decomposition in Columns (3) and (4) are very similar.

As would be expected given the dimensionality of the data, aggregate shocks do not explain much whereas the largest fraction of the dispersion (60%) comes from the volatility due to idiosyncratic shocks affecting the seller-buyer relation. Furthermore, below 10% of the seller-buyer volatility comes from that of the seller per se while the contribution of buyer shocks is twice as large, at 19%. The large contribution of the match-specific shocks is in a way inherent to the estimation strategy, that identifies those components as a residual. For the exact same reason however, we would expect that the impact of the component

[32] Here and in the rest of the paper, the analysis is restricted to variance components calculated on at least 4 points. This restriction can reduce the sample coverage somewhat dramatically, especially when the focus is on highly disaggregated data. See details in Table A.1.

rapidly shrinks once seller-buyer growth rates are aggregated, since they average zero, by definition. As we will show now, this is not what we find using our trade networks data.

4 Volatility in the small

4.1 Theoretical framework

In this section, we use the structure of the data and the estimated sources of volatility to discuss the determinants of volatility *in the small.* Our object of interest is the volatility of a firm's sales:

$$Var(g_{st}) = \frac{1}{T}\sum_t (g_{st} - \bar{g}_s)^2$$

where g_{st} is the growth rate of seller s sales in the destination and $\bar{g}_s$ the mean growth rate of seller s, computed over time. To alleviate notations, the grographic indices are neglected here but we calculate one such variance term for each destination that seller s serves.

At the level of individual firms, the volatility of sales is a weighted average of the variances and covariances of firm-to-firm growth rates, observed in the sub-sample of trade flows involving a single exporter:

$$\begin{aligned} Var(g_{st}) &= Var\left(\sum_{b\in B_s} w^s_{bt-1} g_{sbt}\right) \\ &= Var(f_{Ct}) + Var(f_{st}) + \underbrace{Var\left(\sum_{b\in B_s} w^s_{bt-1}(f_{bt} + \nu_{sbt} + BSIC_{bt})\right)}_{Diversifiable\ components} + Cov \end{aligned} \tag{14}$$

where $w^s_{bt-1} \equiv \frac{x_{sbt-1}}{\sum_{b\in B_s} x_{sbt-1}}$ is the weight of buyer b in seller s (intensive) sales in period $t-1$ and B_s is the set of buyers connected to seller s.[33] The second line uses the decomposition of firm-to-firm components derived before. The Cov term now represents a sum of covariance terms across all types of shocks plus an additional set of spatial covariance terms due to the potential correlation of the buyer-related terms across the different clients belonging to seller s' portfolio. For instance, we shall expect the $BSIC_{bt}$ components to covary positively across buyers connected to the same seller, because they are simultaneously affected by the productivity shocks affecting their common partner.

Equation (14) summarizes the main insight of this section. In presence of multiple sources of volatility, the variance in the small can be thought as the sum of a number of variance and covariance terms, that each depends on one specific source of volatility. Namely, the first term in equation (14) can be interpreted as the aggregate component of the volatility in the small. Since macroeconomic

[33] Appendix A explains how the decomposition extends to extensive trade flows.

shocks are identified under the assumption that their impact is uniform across all exporters selling the same good to a given destination, this component is unlikely to explain why firms have heterogeneous volatilities. The second term in equation (14) is a better candidate to explain such heterogeneity. It represents the micro-level volatility induced by shocks that are specific to the seller. If the variance of these shocks is heterogeneous along the distribution of firms, it can potentially explain the heterogeneity observed in the data. This is the avenue that most of the theoretical literature has followed up to now (Gabaix, 2011, Section 2.5 for instance).

Our framework offers an additional explanation for such an heterogeneity. This additional explanation hinges on the prevalence of buyer and buyer-seller shocks, and the extent to which the structure of the firm's portfolio of customers offers a natural hedging against these shocks. The intuition is better conveyed under three simplifying assumptions, namely that i) the relative share of each buyer in the firm's portfolio is constant over time, ii) different types of shocks are orthogonal to each other, within firm s' portfolio and iii) individual shocks are i.i.d. with a variance that is homogeneous across firms. Under these assumptions, equation (14) simplifies into:

$$Var(g_{st}) = \sigma_C^2 + \sigma_{iS}^2 + (\sigma_{iB}^2 + \sigma_{iSB}^2) \sum_{b \in B_s} w_b^{s\,2} + \phi_s(BSIC_{bt})$$

where σ_C^2, σ_{iS}^2, σ_{iB}^2 and σ_{iSB}^2 respectively refer to the volatilities of the macro component, the seller-specific effect, the buyer-specific term and the match-specific shock. $\phi_s(BSIC_{bt})$ encompasses the variance and covariance terms induced by the buyer-specific input cost term.

This simplified decomposition makes it clear that the prevalence of buyer and buyer-seller shocks as a driver of volatility in the small not only depends on the variance of the shocks, but also on the degree to which they are diversified across customers. Everything else being equal, firms with more diversified sales (with a small Herfindahl index $Herf_s \equiv \sum_{b \in B_s} w_b^{s\,2}$) end up being less volatile because the diversification of their portfolio helps smooth out the impact of shocks to individual buyers. This argument explains that the buyer and buyer-seller specific shocks will be labelled "diversifiable" in what follows.

The presence of these shocks in the analytical framework is also important because it offers a potential source of heterogeneity in individual volatilities which is *endogenous* to the firm. By choosing to diversify its trade network, a firm may benefit from less volatile exports. Such a strategy may be optimal if the firm is risk-averse, if the cost of such diversification is not too high and if the firm cannot hedge using outside instruments (e.g. financial markets). While this paper does not pretend to say anything normative about the extent of diversification, we argue that this dimension is a potential source of volatility in firm-level sales.

Finally, the diversification argument is probably broader than in the limited setting of this particular analysis. In our framework, the firm could diversify against buyer-specific shocks, while also smoothing the impact of aggregate shocks, by selling to more markets. Selling to a broader set of markets is indeed a way for the firm to hedge against country-specific shocks, among which the idiosyncratic shocks to buyers. This possibility is at the root of the argument in Tenreyro et al. (2012), even though they apply it to the volatility *in the large*. Another margin of diversification concerns the product dimension. In presence of product-specific (supply and demand) shocks, a firm would dampen the volatility of her sales by producing a broader portfolio of products. The trade literature has recently emphasized the product margin of international trade (see, among others, Mayer et al. (2014) or Bernard et al. (2011)). However its role on the volatility of sales has not yet been discussed. In what follows, we focus on the buyer margin of diversification because this is where our data are the most valuable. We show that this source of diversification indeed matters for the volatility in the small, since more diversified firms display less volatile sales. Whenever possible, we also control for the degree of cross-product diversification and show that it does matter for the volatility in the small. Extending the exercise to other sources of diversification in a more systematic way could be an interesting avenue of research.

4.2 Empirical results

Table 8 summarizes the results, for the mean firm in the sample. The structure of the table is the same as in Table 7. Columns (1) and (2) respectively report the mean and standard deviation of each variance component, calculated over the population of sellers $\times$ destinations. Column (3) is the contribution of each variance component to the level of volatility. Column (4) is the partial correlation of each variance component with the overall volatility of the firm.

– Table 8 about here –

In comparison with the results for seller-buyer sales (Table 7), the covariance terms are now larger and contribute more to the overall dispersion of volatility measures. In particular, the mean covariance term is now positive, which is consistent with the buyer-specific input cost indices covarying positively between buyers connected to the same seller. Beside the covariances, the most important change concerns the relative contribution of different types of shocks as a source of cross-sectional *dispersion* in volatilities. In particular, the contribution of the "macro" shocks is doubled, but remains negligible. At this level of aggregation, the seller-specific shocks are now a substantial source of volatility and explain almost 40% of the dispersion observed in the cross-section. This result is consistent with models in which the variance of idiosyncratic supply shocks is heterogeneous across firms. Seller-specific shocks are however unable to fully explain the heterogeneity in the data. Buyer and buyer-seller shocks respectively contributes to 20 and 43% of the total heterogeneity in the data. Part of this contribution comes from firms being unequally diversified in the buyer dimension, as we show now.

– Table 9 about here –

Table 9 displays the analysis of the determinants of volatility in the small. Namely, we regress the variance of firm- and destination-specific sales on a set of explanatory variables, including a measure of how diversified the firm is. We use as left-hand side variable either the total variance of sales ($Var(g_{st})$), in columns (1) and (4), or the component attributable to diversifiable shocks ($Var\left(\sum_{b \in B_s} w^s_{bt-1}(f_{bt} + \nu_{sbt})\right)$), in columns (2) and (5), or the component attributable to non-diversifiable shocks ($Var\left(f_{st}\right)$), in columns (3) and (6). Each regression controls for one source of unobserved determinants of volatility using either sector $\times$ destination fixed effects (Columns (1)-(3)) or firm fixed effects (Columns (4)-(6)). For the former specification, the coefficients are identified across firms serving the same destination with the same type of goods. With firm fixed effects, the identification is instead obtained across markets, within firm.

Whatever the structure of fixed effects, results indicate a negative relationship between the variance of a firm's sales and the diversification of its portfolio of customers as measured by the inverse of the Herfindahl index of sales. More diversified firms display significantly less volatile sales. The correlation is especially strong when the diversifiable component of fluctuations is used as left-hand side variable, in columns (2) and (5). The elasticity is then equal to .52 when identified across firms and .45 when the identification is across markets within firms. This represents almost twice the elasticity of the total variance. We interpret the difference in estimated elasticities as evidence that the relationship between a firm's diversification and the volatility of its sales is mostly driven by the impact that diversification has on the firm's exposure to buyer and buyer-seller shocks. Absent such shocks, the significant relationship identified in the data would be difficult to rationalize.[34]

Apart from the Herfindahl of sales, the regressions summarized in Table 9 also control for other drivers of heterogeneity in firm-level fluctuations, notably the size of the firm, the diversification of sales in the product dimension, the experience of the firm in the destination market or proxies for intra-firm trade. The most robust result here is the correlation with the firm's size. Namely, large exporters tend to display less volatile sales, independent of their degree of diversification. This result confirms the correlation found in the previous literature (Davis et al., 2009; Kelly et al., 2013), through with different data and using different dimensions for identification. In contrast, the impact of the firm's experience in the destination is not clear. This is also true of the impact of ownership linkages and diversification in the product dimension.

This closes our analysis of volatility in the small. At the level of individual firms, we have shown that i) individual shocks, especially seller-specific and

[34] By definition, seller-specific shocks cannot be diversified. It is thus surprising that the non-diversifiable component of the variance is still correlated with the Herfindahl of sales in column (3). The correlation however disappears once seller fixed effects are controlled for in column (6). In unreported results, we have also observed that the negative correlation obtained in column (3) is driven by firms which seller-specific component is poorly identified, because those firms are connected to a limited number of buyers. Once those firms are neglected from the analysis, the coefficient on the Herfindahl of sales in column (3) also turns out non-significant.

seller-buyer shocks, generate most of the volatility, ii) these shocks also contribute to explain the heterogeneity in the degree of sales volatility across firms and destination markets, iii) the volatility of sales is correlated with the firm's size and the degree of diversification of its portfolio of customers. We now turn to the analysis of fluctuations *in the large* and ask whether the above-mentioned results disappear when data are further aggregated.

5 Volatility in the large

5.1 Theoretical framework

In the aggregate, the object of interest is the volatility of aggregate bilateral exports, which we define as:

$$Var(g_t) = \frac{1}{T}\sum_t (g_t - \bar{g})^2$$

where g_t is the growth rate of (intensive) aggregate exports to a market and $\bar{g}$ the mean growth rate over the period under consideration. Again, the geographic indices are neglected to simplify notations but the focus is on exports from France to a given European destination.

Using the same logic as in Section 4, the variance of aggregate sales decomposes into the structural drivers identified in Section 3:

$$\begin{aligned} Var(g_t) &= Var\left(\sum_{s\in S}\sum_{b\in B_s} w_{st-1} w^s_{bt-1} g_{sbt}\right) \\ &= Var(f_{Ct}) + Var\left(\sum_{s\in S} w_{st-1} f_{st}\right) + Var\left(\sum_{b\in \bigcup_{s\in S} B_s} w_{bt-1} f_{bt}\right) \\ &+ Var\left(\sum_{s\in S}\sum_{b\in B_s} w_{sbt-1}\nu_{sbt}\right) + Var\left(\sum_{b\in \bigcup_{s\in S} B_s} w_{bt-1} BSIC_{bt}\right) + Cov \end{aligned} \tag{15}$$

where w_{st-1}, w_{bt-1} and w_{sbt-1} denote respectively the shares of seller s, buyer b and the pair (s, b) in the value of the aggregate (intensive) trade flow. Cov is a set of covariance terms between the different shocks involved in (8).[35]

Equation (15) is the counterpart to equation (14), for the volatility in the large. Again, it is useful to study its implications under the simplifying assumptions that i) the relative shares are constant over time, ii) different types of shocks are orthogonal to each other and iii) individual shocks are i.i.d. and

[35] In equation (15), Cov comprises the covariance terms involving the different types of shocks. It also encompasses any spatial correlation among the individual shocks affecting different nodes in the network.

equally volatile across individuals. Under these assumptions, equation (15) simplifies into:

$$Var(g_t) = \sigma^2_C + \sigma^2_{iS} \sum_{s \in S} w^2_s + \sigma^2_{iB} \sum_{b \in \bigcup_{s \in S} B_s} w^2_b + \sigma^2_{iSB} \sum_{s \in S} \sum_{b \in B_s} w^2_{sb} + \phi(BSIC_{bt})$$

This decomposition helps emphasize the different sources of "granular" fluctuations in the data. Apart from the macroeconomic volatility (the σ^2_C term in the above equation), aggregate fluctuations arise from three sources of individual shocks, the seller-specific, the buyer-specific and the seller-buyer components. Their relative contribution to the volatility in the large depends on the variance of the shocks *and* the extent to which they are diversified in the aggregate economy. The argument is strictly the same as in Gabaix (2011). In presence of idiosyncratic supply-shocks (i.e. for $\sigma^2_{iS} \neq 0$), the concentration of sales along the distribution of firms is positively related with the variance in the aggregate because more concentration implies that shocks to the largest sellers are not compensated by shocks to smaller firms. As illustrated in Figure 1, first panel, the distribution of individual sales is extremely concentrated, the Herfindahl index of sales across sellers $Herf_S \equiv \sum_{s \in S} w^2_s$ is large and thus idiosyncratic shocks to sellers can deliver granular fluctuations.

What our analysis shows is that the argument naturally extends to the concentration of sales across buyers (the inverse of $Herf_B \equiv \sum_{b \in \bigcup_{s \in S} B_s} w^2_b$) and the concentration of transactions across seller-buyer pairs (the inverse of $Herf_{SB} \equiv \sum_{s \in S} \sum_{b \in B_s} w^2_{sb}$). Again, the prevalence of those shocks as a source of aggregate fluctuations depends on their variance as well as the concentration of trade. Since both Herfindahl indices are large in our trade data (Figure 1, second and third panels), we expect the buyer-specific and the seller-buyer shocks to matter for aggregate fluctuations.

5.2 Empirical results

Table 10 gives summary statistics on the variance in the large, and its components. The structure is the same as in Table 8 but the number of observations is now reduced to eleven points, which implies that the statistics are more strongly affected by outliers. We thus also report the exact decomposition obtained for each country in Figure 4.

– Table 10 about here –

– Figure 4 about here –

Once the analysis focuses on aggregate fluctuations, the contribution of aggregate shocks naturally increases. Namely, 6% of the variance in the median destination is attributable to macroeconomic or sectoral shocks (Table 10, Column (3)), the contribution varying between 3% in Finland and 16% in Belgium. Even though this contribution is now larger, it is still small in comparison to the combined effect of all individual shocks. This comes from the huge concentration of sales along the distribution of trade flows, the granularity of the data.

Within the set of individual shocks, all three types of shocks contribute substantially to the overall variance (see Table 10, Column (3)), the extent to which it is the case varying across countries. For example, seller shocks are the most important source of granular fluctuations for Spain. In Denmark, it is shocks to individual buyers which matter the most while seller-buyer shocks are key to explaining the variance of sales towards Finland. Across countries, seller-buyer shocks explain a substantial share of the differences in volatilities while the role of buyer and seller shocks is smaller, though still significant. The relative contribution of different types of shocks is in fact remarkably similar to its counterpart observed in the small (see the comparison of Tables 8 and 10). Following Kelly et al. (2013), one would expect that the impact of buyer and seller-buyer shocks is reduced in the aggregate because large firms tend to diversify customer-related risks better. On average, this is true in our data, however with some non-linearities. In particular, we observed in section 2.2 that the largest exporters are in fact poorly diversified. Since very large firms are those that matter in the aggregate, this explains that the impact of buyer and seller-buyer shocks remains significant, even in the aggregate.

In our framework, the degree to which different individual shocks induce fluctuations in the large is due to the interplay between the variance of these shocks and the concentration of sales across individuals, the amount of granularity. Were shocks equally volatile and orthogonal across firms, the aggregate effect of seller-specific shocks (resp. buyer-specific and seller-buyer shocks) would be just proportional to the Herfindahl of sales, across sellers (resp. across buyers and seller-buyer pairs). If instead the volatility of seller-specific shocks (resp. buyer-specific and seller-buyer shocks) were inversely proportional to sellers' size (resp. buyer-specific and seller-buyer size), then we would observe no correlation between diversification and aggregate volatility. To assess the importance of diversification on aggregate volatility, Figure 5 plots the variance attributable to the direct effect of each family of individual shocks against the Herfindahl of sales, computed across the corresponding individual trade flows. For all three types of shocks, the correlation is strongly positive.

– Figure 5 about here –

Focusing first on the upper-left panel, regarding the impact of seller-specific shocks, we see that the correlation is positive and significant at 1%, in agreement with the intuition. The seller-specific component of aggregate fluctuations is large partly because sales are highly concentrated across sellers. This is especially true in countries like Spain, which contributes to the large volatility of sales towards this destination. The correlation between sales concentration along the buyer-seller dimension and the variance induced by match-specific shocks is also strongly positive and significant (bottom-left panel). Instead, the coefficient of regression is reduced and found non-significant when the focus is put on the buyer-specific components of aggregate fluctuations. Despite substantial concentration of purchases and heterogeneity across destinations, the diversification of sales along the buyer dimension does not induce much heterogeneity in aggregate volatility.

– Figure 6 about here –

Figure 6 illustrates why this is the case. Intuitively, the reason why the concentration of sales matters for aggregate fluctuations is because shocks affecting the largest individuals in the network cannot be compensated by shocks to smaller firms. However, if these individuals happen to display less volatility than smaller ones, the forces toward granularity are reduced. Figure 6 shows that this is indeed the case. Namely, seller-specific, buyer-specific and seller-buyer shocks affecting the top 1% of exporters, importers, and seller-buyer pairs, respectively, tend to be less volatile than in the rest of the distribution. This moderates the forces towards granularity. This is especially true for buyer-specific shocks, which are almost three times less volatile for buyers in the upper tail of the distribution than for smaller buyers. This explains that the impact of more concentration in purchases does not trigger much additional volatility.[36] One potential interpretation of this result is that large importers are already able to "aggregate" risk, say because they intermediate the demand of other local customers. This reduces the volatility of their demand. The argument is consistent with outside evidence on the role of wholesalers in international trade (Bernard et al., 2010).

To summarize, these results show that the prevalence of individual shocks is still substantial in aggregate data, because aggregate trade flows are strongly concentrated along the distribution of trade partners. As a consequence, shocks affecting the largest exporting firms and/or the largest seller-buyer transactions do not compensate with shocks to smaller trade flows, thus generating substantial aggregate fluctuations. In comparative terms, the role of buyer-specific shocks is somewhat reduced, because the strong concentration of import purchases is compensated by larger importers displaying less risk.

6 Conclusion

In this paper, we provide a forensic account of the origin of fluctuations in sales at the level of individual firms as well as in the aggregate. We first propose a new methodology for identifying different categories of shocks in disaggregated growth data. We next show that individual shocks together with the structure

[36]While Figure 6 summarizes the results, we also computed the correlation coefficients between individual variances and the relevant weights. We found: i) a correlation of -0.0077 between the variance of seller-specific shocks and w_s^2, ii) a correlation of -0.0115 between the variance of buyer-specific shocks and w_b^2 (significant at 1%), and iii) a correlation of -0.0034 between the variance of the seller-buyer shocks and w_{sb}^2. The larger degree of correlation between the variance of buyer-specific shocks and the weight of individual importers is consistent with the above story.

of trade networks help explain the volatility of sales and their heterogeneity across firms and markets. In the small, shocks related to customers are an important component of volatility. Differences in the structure of firms' portfolio of buyers are key to account for their differences in volatility. In the large, individual shocks are shown to be the main force behind aggregate volatility. The differences in the volatility of French exports across countries are tightly linked to the interplay between the sources of individual shocks and the structure of trade networks.

Our model is highly stylized. Developing a model of the dynamics of seller-buyer trade relationships and their aggregate implications is a natural extension of the present paper. Such model may help better understanding the extensive margin of individual sales. Indeed, the entry and exit of buyers in sellers' portfolios is found to have a non negligible impact on sellers' dynamics. We plan to explore this dimension further in future work.

A Role of extensive adjustments

A.1 The intensive and extensive margins of export growth

At the level of individual firms, the growth rate of destination-specific sales is defined as:

$$g_{st}^{Tot} = \ln x_{st} - \ln x_{st-1} = \ln\left(\sum_{b \in B_{st}} x_{sbt}\right) - \ln\left(\sum_{b \in B_{st-1}} x_{sbt-1}\right)$$

where x_{sbt} is the value of exports from seller s to buyer b at date t and B_{st} the set of buyers in seller s' portfolio at date t.

This decomposes into an intensive and an extensive components as follows:

$$g_{st}^{Tot} = g_{st} + g_{st}^{Ext.} \tag{A.1}$$

The intensive component

$$g_{st} = \ln\left(\frac{\sum_{b \in B_s} x_{sbt}}{\sum_{b \in B_s} x_{sbt-1}}\right)$$

is driven by changes in sales to buyers active in the firm's portfolio at dates t and $t-1$ ($d \ln x_{sbt}$ for $b \in B_s$ and $B_s \equiv B_{st} \cap B_{st-1}$ the set of incumbent buyers in seller s portfolio). This is the growth component which the decomposition in Section 3 studies. The extensive component is defined as

$$g_{st}^{Ext.} = \ln\left(\frac{\sum_{b \in B_{st}} x_{sbt}}{\sum_{b \in B_s} x_{sbt}} \frac{\sum_{b \in B_s} x_{sbt-1}}{\sum_{b \in B_{st-1}} x_{sbt-1}}\right)$$

It thus measures the contribution to sales growth of new entrants, in relative terms with respect to the contribution of buyers that have stopped importing from s between $t-1$ and t.

In the aggregate, the growth of exports decomposes as follows:

$$g_t^{Tot} = g_t + g_t^{Ext-buyer} + g_t^{Ext-seller} \tag{A.2}$$

g_t^{Tot} represents the growth of aggregate exports to a destination:

$$\begin{aligned} g_t^{Tot} &= \ln x_t - \ln x_{t-1} \\ &= \ln\left(\sum_{s\in S_t} x_{st}\right) - \ln\left(\sum_{s\in S_{t-1}} x_{st-1}\right) \\ &= \ln\left(\sum_{s\in S_t}\sum_{b\in B_{st}} x_{sbt}\right) - \ln\left(\sum_{s\in S_{t-1}}\sum_{b\in B_{st-1}} x_{sbt-1}\right) \end{aligned}$$

where S_t is the set of sellers serving the destination at time t.

The intensive component is

$$g_t = \ln\left(\frac{\sum_{s\in S}\sum_{b\in B_s} x_{sbt}}{\sum_{s\in S}\sum_{b\in B_s} x_{sbt-1}}\right)$$

It is driven by changes in the sales of seller-buyer transactions present at dates t and $t-1$ (the set $(s,b) \in \bigcup_{s\in S} B_s$), which itself is defined on the subset of incumbent exporters $S = S_t \cap S_{t-1}$. This component of growth is the object of analysis in Section 4.

At the aggregate level, the extensive margin can be decomposed into a buyer and a seller components. The buyer component of the extensive margin is defined as

$$g_t^{Ext-buyer} = \ln\left(\frac{\sum_{s\in S}\sum_{b\in B_{st}} x_{sbt}}{\sum_{s\in S}\sum_{b\in B_s} x_{sbt}} \times \frac{\sum_{s\in S}\sum_{b\in B_s} x_{sbt-1}}{\sum_{s\in S}\sum_{b\in B_{st-1}} x_{sbt-1}}\right)$$

It represents the weight of new buyers in total sales of incumbent sellers, in relative terms with respect to the weight of purchases by buyers that exit the portfolio between $t-1$ and t. The seller component of the extensive margin is in turn

$$g_t^{Ext-seller} = \ln\left(\frac{\sum_{s\in S_t} x_{st}}{\sum_{s\in S} x_{st}} \times \frac{\sum_{s\in S} x_{st-1}}{\sum_{s\in S_{t-1}} x_{st-1}}\right)$$

$g_t^{Ext-seller}$ thus measures the weight of new sellers in total exports relative to the weight of sellers that exited the market.

The analysis in the main body of the text focuses on fluctuations in the intensive components of g_{st}^{Tot} and g_t^{Tot}. This is motivated by evidence in Table 4 that intensive flows are the most important source of growth in our data. We now discuss the extent to which the neglected extensive adjustments further amplify fluctuations in the small and in the large.

A.2 Volatility in the small and the extensive margin

Using equation (A.1), the overall volatility of firm-level sales decomposes as follows:

$$Var(g_{st}^{Tot}) = Var(g_{st}) + Var(g_{st}^{Ext.}) + 2Cov(g_{st}, g_{st}^{Ext.}) \tag{A.3}$$

While we focus the analysis on the $Var(g_{st})$ component, adjustments at the (buyer) extensive margin might contribute to generating fluctuations in firm-specific sales. This is especially likely to be the case if extensive adjustments correlate positively with fluctuations at the intensive margin.[37] The extent to which it is indeed the case is an empirical question which Table A.2 intends to solve.

– Table A.2 about here –

For the median firm, the intensive component of the variance represents 92% of the overall variance (Table A.2, Column (3)). Contrary to expectations, the covariance between the intensive and extensive components is negative, on average. This contributes to reducing the overall variance. However, the magnitude of this term substantially varies across firms which precludes any strong interpretation. While the intensive margin is the most important source of volatility, results in the fourth column of Table 3 show that both the intensive and the extensive margins contribute to the *dispersion of volatilities across firms.*[38]

A.3 Volatility in the large and the extensive margin

Using equation (A.2), the overall volatility of aggregate sales in turn decomposes as follows:

$$Var(g_t^{Tot}) = Var(g_t) + Var(g_t^{Ext-buyer}) + Var(g_t^{Ext-seller}) + Cov \tag{A.4}$$

where Cov now includes all covariance terms involving one of the three components of (A.2).

Adjustments at the buyer or seller extensive margin might contribute to generating fluctuations in aggregate sales. Table A.3 quantifies the extent to which it is the case.

– Table A.3 about here –

At the aggregate level, the buyer extensive margin is clearly a negligible source of fluctuations. Its variance is thirty times smaller than the overall variance of export growth and it contributes to a tiny share of the dispersion in

[37]Note that this is likely to be the case in a dynamic model with a fixed cost of serving a buyer. In such model a negative productivity shock to a seller would reduce sales to each of its partners, and eventually force it to stop serving some of these buyers, if the operational profits their demand generates is not sufficient to cover the fixed cost.

[38]These results are obtained in the sub-sample of 52,831 seller-destination pairs for which all variance components can be identified over at least 4 years. If one considers instead all the firm×destinations in our dataset, we find that the intensive margin contributes to about 2/3 of the crossectional dispersion in volatilities.

volatilities across countries. The seller extensive margin is quantitatively more important, but still small in comparison with the intensive component of fluctuations. In aggregate data even more than at the level of individual firms, the intensive margin is the main source of fluctuations in trade growth and the most important driver of heterogeneity in volatilities, across destinations.

B Details on the estimation strategy

The estimated equation takes the following form:

$$\tilde{g}_{sbt} = (1+\lambda) f_{Ct} + f_{st} + (1+\lambda) f_{bt} + \nu_{sbt}$$

We follow Abowd et al. (1999) and assume:

$$E[\nu_{sbt}|(s,t);(b,t)] = 0$$

In words, this condition restates the exogeneity condition for our specific case. The residual ν_{sbt} is orthogonal to the buyer$\times$time and the seller$\times$time effects, conditional on the other effects. Three things are worthy of note. First, the condition holds at every time period. Second, the growth measure can only be computed for relations between a seller and a buyer that last between $t-1$ and t (we discuss the robustness of this assumption later). Third, even though a buyer's identity is country-specific, this is not the case for the sellers since they may sell in all countries. Hence, the assumption holds across all observations of a given seller to his buyers in the 11 countries in the data. This last remark is important in view of our discussion of the so-called "limited mobility bias" discussed below. Estimation of this model is simple and has been widely discussed in the literature, starting with Abowd et al. (2002) who were the first to provide the full identification conditions of the fixed effects to be estimated.

Estimation of λ: Estimating the above equation using Abowd et al. (1999) requires that we first estimate the λ parameter. As explained in the text, we identify the parameter using an additional orthogonality condition suggested by the theoretical model, namely equation (12). Under the true value of λ, the model tells us that the seller and buyer components should be orthogonal to each other. For any $\lambda' \neq \lambda$, we have instead:

$$\begin{aligned} Cov(f_{st}^{\lambda'}, f_{bt}^{\lambda'}) &= Cov\left(f_{st}, (1+\lambda) f_{bt} + (\lambda' - \lambda) g_{bt}\right) \\ &= (\sigma - 1)(\eta - 1)(\lambda' - \lambda) w_{st-1} Var(\varepsilon_{st}) \end{aligned}$$

where $f_{st}^{\lambda'}$ and $f_{bt}^{\lambda'}$ denote the seller and buyer components of an equation using as left hand side variable $\tilde{g}_{sbt}^{\lambda'} \equiv g_{sbt} + \lambda' g_{bt}$. Misspecifying the LHS variable of equation (9) thus augments the buyer-specific component with an additional term which is systematically correlated with the (theoretical) seller-specific effect. This shall induce a covariance between the estimated seller and buyer effects. The algorithm implemented to estimate λ uses this prediction of the model and selects the value for λ which satisfies the orthogonality condition implied by the model. Note that the algorithm is straightforward to implement since the value of the covariance is monotonous in $(\lambda' - \lambda)$ for all values of the structural parameters such that $\sigma > 1$ and $\eta > 1$: Any value of $\lambda' < \lambda$ (resp. $\lambda' > \lambda$) implies a negative (resp. positive) covariance between the estimated seller and buyer components.

Limited Mobility Bias: Abowd et al. (2004) were the first to note that, in models with two-way effects, even when data were simulated with no correlation between the individuals at each side of the graph (here, between buyers and sellers), estimating these effects and then computing the correlation between the resulting effects yielded a negative correlation. This finding has been found multiple times in various types of data sources for which these two-way effects were relevant modeling tools. The intuition for this result is quite straightforward. In such additive models, when an estimation error is made on one effect, there is a corresponding estimation error of the opposite sign on the other effect. Because the standard error of these effects decreases as the number of observations used to estimate them increases, the larger the number of buyers connected to a seller, or conversely the number of sellers connected to a buyer, the more precise these effects become.[39] However, the complex nature of the bipartite graph used with such data makes the analysis involved. In a very useful paper, (Andrews et al., 2008) have investigated this question more systematically. And, indeed, they confirm that the bias is larger the smaller the number of movers, or, in our application the less connected our buyers and sellers are.

It is possible to compute, but only for simple examples, the bias formula (see (Andrews et al., 2008) page 682, eq. (27)). Assume that there are N^S sellers, labeled 1 to N^S and N^B buyers. Assume that seller 1 sells one good to buyer 1 who also buys one good from seller 2, seller 2 sells one good to buyer 2 who also buys one good from seller 3.... until N^S sells to $N^B = N^S$ who buys from seller 1. Now, let us think about the equivalent situation with M times more buyers and the same number of sellers, in which seller 1 sells to M buyers, labeled 1 to M, who all buy from seller 2.... and so on until the circle is completed with our N^S sellers and our $N^B = M * N^S$ buyers. Then, it is possible to prove that the correlation between the buyer and the seller effects is

$$-\frac{\sigma_\nu^2}{N^*}\left(\frac{k}{M} - N^S\right)$$

where σ_ν^2 is the variance of the residual, k is a constant such that $k > M \times N^S$, and N^* is the total number of observations which grows with M. Andrews et al. (2008) follow on Abowd et al. (2004) by using simulation to assess the magnitude of the bias in the correlation between the individual effects in more general setups. They confirm how connectedness is directly related to the magnitude of this bias.

[39] Since such models were first applied to workers and firms, the more workers moved between firms the more precise the estimates, hence the choice of the "limited mobility" name.

We also adopt this strategy in the following, arguing that the structure of the network by itself induces a bias in the estimated seller and buyer effects. To quantify the magnitude of this bias, we generate uncorrelated seller and buyer effects from a normal distribution with fixed, known variance for each node of the network, as well as a residual, also drawn in a normal distribution.[40] Adding these effects, we generate simulated growth rates. These growth rates are used to estimate the seller and buyer effects using the AKM procedure and, then, compute the associated correlation between the two. This procedure is repeated 100 times. This yields a distribution of the bias using our simulated effects and the realized structure of the network since, by construction, the *true* correlation between these effects is equal to zero. We select the mean of this distribution as our target bias, which is -0.0655 in our data. We then take into account the limited mobility bias by targeting this value for $Cov(f_{st}, (1+\lambda)f_{bt})$ instead of the strict orthogonality condition (12). **** Do we say a word on the correlation between the true shocks and the estimated shocks in our simulation? ****

Survival bias: It is quite straightforward to adapt our methodology to take into account the survival bias embedded in our computation of the growth rate of the exports of a seller to a buyer. Namely, the use of growth rates as LHS variable implies that we de facto neglect all combinations of shocks which destroy the relationship, either because the seller dies, or because it is the buyer which exits the market, or simply because both nodes stay active but no longer trade together. Since such combinations of shocks are probably not randomly drawn in the distribution of all possible combinations, neglecting those observations might induce a bias. In J.M. et al. (2001), it is shown that a valid procedure for the type of data at hand consists in weighting each observation by the inverse of the death probability of the observation (hence, of the trade relationship). The main problem in implementing this approach with the data at hand is that we do not know much about sellers (in terms of observables), not to mention buyers for which we know close to nothing except the products they buy, their past purchases, and the country in which they operate.

[40] For all three components, the variance of the underlying normal distribution is calibrated using the mean variance estimated when equation (9) is estimated assuming $\lambda = 0$.

Table 1: Summary statistics on trade networks

	Value of exports (bil.)	# French sellers	# foreign buyers	# pairs of buyer-seller
Belgium	26.6	29,941	74,427	225,823
Denmark	2.8	8,567	9,248	22,008
Finland	1.85	5,420	5,379	12,243
Germany	50.2	25,078	122,568	249,197
Ireland	2.54	6,508	6,857	16,804
Italy	32.0	20,565	100,115	192,628
Netherlands	15.5	16,851	35,080	73,568
Portugal	4.59	11,980	20,331	44,957
Spain	35.5	22,038	80,178	166,738
Sweden	5.08	7,896	10,757	21,832
United Kingdom	30.6	19,289	52,596	115,992
EU11	207	42,888	334,905	1,141,326

Notes: Summary statistics computed on 2007 data describing French bilateral exports. The last line corresponds to the 11 members of the European Union pooled together. The table does not include the transactions for which the CN8 product code is not reported (19,803 sellers accounting for less than 0.05% of exports). Column (1) reports the value of the aggregate trade flow, in billions euros. Columns (2)-(4) respectively report the number of sellers, buyers, and seller-buyer pairs involved in this aggregate trade flow.

Table 2: Concentration of trade flows, by destination

		Concentration across sellers				Concentration across buyers				Concentration ac. seller-buyer pairs			
	Export value (1)	Herf. ac. sellers (2)	Herf. × # sellers (3)	Top 10% (4)	10 largest (5)	Herf. ac. buyers (6)	Herf. × # buyers (7)	Top 10% (8)	10 largest (9)	Herf. ac. pairs (10)	Herf. × # pairs (11)	Top 10% (12)	10 largest (13)
Belgium	26.6	0.007	196	88%	19%	0.007	493	97%	20%	0.005	1178	94%	18%
Germany	50.2	0.003	83	90%	12%	0.003	358	97%	12%	0.002	455	95%	9%
Denmark	2.76	0.009	79	87%	19%	0.010	93	90%	20%	0.009	188	90%	18%
Spain	35.5	0.018	404	89%	27%	0.013	1026	96%	29%	0.007	1219	94%	22%
Finland	1.85	0.011	57	87%	24%	0.012	62	91%	25%	0.009	116	90%	23%
UK	30.6	0.009	179	90%	20%	0.007	368	96%	21%	0.006	661	94%	17%
Ireland	2.54	0.030	198	90%	40%	0.031	210	93%	39%	0.030	497	93%	39%
Italy	32.0	0.005	107	90%	16%	0.004	365	95%	15%	0.003	614	93%	14%
Netherlands	15.5	0.009	144	90%	24%	0.006	225	95%	21%	0.005	384	94%	19%
Portugal	4.59	0.015	177	87%	24%	0.009	187	93%	24%	0.007	330	92%	22%
Sweden	5.08	0.024	191	90%	32%	0.028	302	94%	37%	0.023	496	93%	29%
EU11	207	0.005	234	90%	15%	0.001	497	94%	7%	0.001	755	94%	6%

Notes: Summary statistics computed on 2007 data describing French bilateral exports. The first column is the value of aggregate exports, in billion euros. Column (2) is the Herfindahl index across exporters, computed as $Herf_S = \sum_{s \in S} w_s^2$ where w_s is the share of exporter s in the total bilateral flow. Column (3) rescales this number by the number one would expect from a uniform distribution of exporters (i.e. $1/S$). Columns (4) and (5) report the share of aggregate exports that is attributable to the top 10% and the largest 10 exporters. Column (6) is the Herfindahl index across importers, computed as $Herf_B = \sum_{b \in B} w_b^2$ where w_b is the share of importer b in the total bilateral flow. Column (7) rescales this number by the number one would expect from a uniform distribution of importers (i.e. $1/B$). Columns (8) and (9) report the share of aggregate exports that is attributable to the top 10% and the largest 10 importers. Column (10) is the Herfindahl index across exporter-importer pairs, computed as $Herf_{SB} = \sum_{(s,b) \in N} w_{sb}^2$ where w_{sb} is the share of the pair in the total bilateral flow. Column (11) rescales this number by the number one would expect from a uniform distribution of pairs (i.e. $1/N$). Columns (12) and (13) report the share of aggregate exports that is attributable to the top 10% and the largest 10 pairs. The last line corresponds to the 11 members of the European Union pooled together.

Table 3: Determinants of firm-level diversification within a country

	ln # **buyers**			ln **Herfindahl**		
	(1)	(2)	(3)	(4)	(5)	(6)
ln value of exports	0.22***	0.21***	0.28***	-0.08***	-0.10***	-0.13***
	(0.022)	(0.010)	(0.015)	(0.013)	(0.006)	(0.010)
(ln value of exports)^2	-0.01***	-0.01***	-0.01***	0.01***	0.01***	0.01***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.000)	(0.000)
ln experience in dest.	0.11***	0.34***	0.13***	-0.06***	-0.22***	-0.10***
	(0.008)	(0.020)	(0.019)	(0.005)	(0.012)	(0.013)
ln # products	0.40***	0.74***	0.53***			
	(0.013)	(0.020)	(0.023)			
ln Herfindahl ac. products				0.27***	0.39***	0.35***
				(0.010)	(0.014)	(0.014)
$\mathbb{1} = 1$ if HQ in dest.	-0.19***	-0.01	-0.02	0.16***	0.02	0.04***
	(0.033)	(0.032)	(0.024)	(0.018)	(0.015)	(0.014)
$\mathbb{1} = 1$ if affiliates in dest.	-0.19***	-0.04	-0.18***	0.13***	0.03	0.13***
	(0.052)	(0.086)	(0.060)	(0.034)	(0.051)	(0.040)
ln potential # of buyers	0.04***	0.00	0.00			
	(0.006)	(0.003)	(0.004)			
ln potential Herfindahl				0.03***	0.09***	0.03***
				(0.004)	(0.014)	(0.006)
FE $Sect \times dest.$	Yes	No	Yes	Yes	No	Yes
FE $Firm$	No	Yes	Yes	No	Yes	Yes
# obs.	158,239	158,239	158,239	158,239	158,239	158,239
R^2	0.184	0.294	0.676	0.100	0.139	0.556

Notes: Standard errors in parentheses clustered in the *destination* × *sector* dimension with ***, ** and * respectively denoting significance at the 1, 5 and 10% levels. "ln potential # of buyers" is the log of a (weighted) average of the number of firms buying at least one variety (whatever the exporter buying it) in each *nc*8 sector in which the exporter is active. "ln potential Herfindahl" is the log of the Herfindahl that the firm would display if it was serving each potential buyer of its *nc*8 products in proportion of their total purchases.

Table 4: Contribution of the intensive and extensive margins to export growth

	Contribution to			
	Individual growth		Aggregate growth	
	Mean	Median	Mean	Median
	g_{st}	g_{st}	g_t	g_t
	(1)	(2)	(3)	(4)
Intensive	0.791	1.000	0.654	0.645
Extensive	0.209	0.000	0.346	0.273
of which:				
Buyer margin			-0.046	0.153
Seller margin	0.209	0.000	0.392	0.120
# of obs.	1,570,494		132	

Notes: Statistics on the decomposition of the total growth into the intensive and extensive margins. The formula are detailed in Appendix A, equations (A.1) and (A.2). Columns (1) and (2) decompose firm-level destination-specific growth rates while Columns (3) and (4) decompose the growth of aggregate bilateral sales. Growth rates are computed annually on the period 1996-2007. The first and third columns give the mean contribution computed on the corresponding sample of yearly growth rates. The second and fourth columns give the median contributions.

Table 5: Summary statistics on the estimated effects

	(1) Mean	(2) Std.Dev	(3) Dimension	(4) Contrib.	(5) Partial Corr.
Firm-to-firm growth g_{sbt}	-0.0140	0.6891	3,184,084		
Macro component f_{Ct}	-0.0539	0.0165	3,949	0.021	0.002***
Seller-specific component f_{st}	0.0000	0.2654	219,354	0.073	0.115***
Buyer-specific component f_{bt}	0.0000	0.3609	798,612	0.223	0.253***
Match-specific residual ν_{sbt}	-0.0000	0.5416	3,184,084	0.631	0.618***
Buyer-specific input cost $BSIC_{bt}$	0.0400	0.1431	798,612	0.006	0.013***

Notes: This table gives the mean (column (1)) and standard deviation (column (2)) of each of the component of seller-buyer growth rates, over the population of estimated effects. The number of estimated effects is displayed in column (3). Column (4) is the median contribution of each growth component to the seller-buyer growth (eg. $Med(f_{st}/g_{sbt})$). The last column is the regression coefficient of each component on the firm-to-firm growth rate. *** indicates significance at the 1% level.

Table 6: Correlation matrix of the estimated growth components

	(1) g_{sbt}	(2) f_t	(3) f_{st}	(4) f_{bt}	(5) ν_{sbt}	(6) $BSIC_{bt}$
g_{sbt}	1.0000					
f_{Ct}	0.0207	1.0000				
f_{st}	0.2974	0.0000	1.0000			
f_{bt}	0.4825	0.0000	-0.0662	1.0000		
ν_{sbt}	0.7861	-0.0000	0.0000	0.0000	1.0000	
$BSIC_{bt}$	0.0689	-0.0157	-0.2560	-0.0758	0.0000	1.0000

Notes: This table gives the correlation matrix between the growth components, in the panel of firm-to-firm growth rates.

Table 7: Summary statistics on the variance components at the firm-to-firm level

	(1) Mean	(2) Std.Dev	(3) Contrib.	(4) Partial Corr.
Variance of				
Firm-to-firm growth g_{sbt}	0.4447	0.3469		
Macro component f_{Ct}	0.0003	0.0000	0.001	0.000***
Seller-specific component f_{st}	0.0585	0.0956	0.090	0.093***
Buyer-specific component f_{bt}	0.1096	0.1228	0.221	0.190***
Match-specific residual ν_{sbt}	0.2892	0.2515	0.681	0.593***
Buyer-specific input cost $BSIC_{bt}$	0.0179	0.0218	0.033	0.029***
Covariance terms Cov	-0.0305	0.2286	-0.110	0.094***
Count observations	273,142			

Notes: This table gives summary statistics on the variance of firm-to-firm growth and its components. The decomposition is based on equation (13), where the "Cov" term is the sum of all covariance terms involving f_{Ct}, f_{st}, f_{bt}, ν_{sbt} and $BSIC_{bt}$. Column (1) reports the mean variance in the population of firm-to-firm relationships, Column (2) its standard deviation. Column (3) is the median contribution of each variance component to the total variance (eg. $Med(var(f_{st})/var(g_{sbt}))$). Column (4) is the partial correlation between each variance component and the overall variance. The sample is restricted to variances computed on at least four growth rates. *** indicates significance at the 1% level.

Table 8: Summary statistics on the variance components, at the firm level

	(1) Mean	(2) Std.Dev	(3) Contrib.	(4) Partial Corr.
Variance of				
Firm growth g_{st}	0.2219	0.1920		
Macro component f_{Ct}	0.0003	0.0001	0.002	0.001***
Seller-specific component f_{st}	0.1087	0.1482	0.413	0.390***
Buyer-specific component f_{bt}	0.0565	0.0659	0.233	0.196***
Match-specific residual ν_{sbt}	0.1120	0.1117	0.523	0.429***
Buyer-specific input cost $BSIC_{bt}$	0.0127	0.0183	0.046	0.041***
Covariance terms Cov	0.2219	0.1920	-0.335	-0.057***
Count observations	52,831			

Notes: This table gives summary statistics on the variance of firm growth and its components. The decomposition is based on equation (14), where the "Cov" term is the sum of all covariance terms involving f_{Ct}, f_{st}, f_{bt}, ν_{sbt} and $BSIC_{bt}$, both within and across buyers belonging to the same firm's portfolio. Column (1) reports the mean variance in the population of firms, Column (2) its standard deviation. Column (3) is the median contribution of each variance component to the total variance (eg. $Med(var(f_{st})/var(g_{st}))$). Column (4) is the partial correlation between each variance component and the overall variance. The sample is restricted to variances computed on at least four growth rates. *** indicates significance at the 1% level.

Table 9: Determinants of the volatility of sales at the firm level

	ln variance of sales					
	Total	Div.	Non-Div.	Total	Div.	Non-Div.
	(1)	(2)	(3)	(4)	(5)	(6)
ln Herfindahl	0.33***	0.52***	0.27***	0.29***	0.45***	0.01
	(0.007)	(0.007)	(0.008)	(0.009)	(0.009)	(0.006)
ln Herf. ac. prod.	-0.03***	-0.03***	0.07***	0.08***	0.09***	0.02*
	(0.007)	(0.007)	(0.008)	(0.014)	(0.015)	(0.009)
ln value of exports	-0.09***	-0.06***	-0.03***	-0.14***	-0.11***	-0.01***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
ln # years	-0.26***	0.10***	-0.49***	-0.14***	-0.22***	0.59***
	(0.011)	(0.011)	(0.013)	(0.016)	(0.016)	(0.010)
Entrant	0.05***	-0.11***	0.16***	0.06**	-0.03	0.25***
	(0.012)	(0.012)	(0.014)	(0.026)	(0.026)	(0.016)
Young exporter	0.01	-0.02**	0.10***	-0.00	0.02	-0.02**
	(0.010)	(0.010)	(0.012)	(0.018)	(0.018)	(0.011)
$\mathbb{1} = 1$ if HQ in dest.	0.03	0.12***	0.15***	0.10***	0.14***	0.05***
	(0.025)	(0.025)	(0.029)	(0.030)	(0.030)	(0.019)
$\mathbb{1} = 1$ if aff. in dest.	-0.05	-0.02	-0.19***	-0.00	-0.09	-0.01
	(0.043)	(0.043)	(0.051)	(0.058)	(0.059)	(0.037)
FE $Sect \times dest.$	Yes	Yes	Yes	No	No	No
FE $Firm$	No	No	No	Yes	Yes	Yes
Observations	52,831	52,831	52,831	52,831	52,831	52,831
Adjusted R-squared	0.110	0.147	0.143	0.315	0.326	0.805

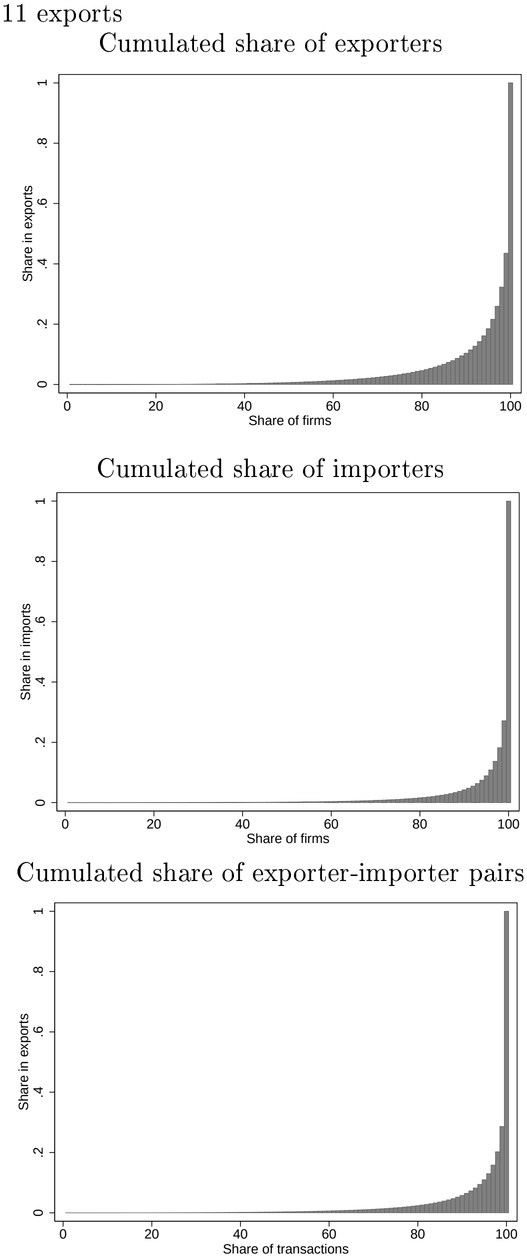
Notes: Robust standard errors in parentheses with ***, ** and * respectively denoting significance at the 1, 5 and 10% levels. "Total", "Div." and "Non-Div." respectively denote the total variance of sales ($Var(g_{st})$), the variance of the diversifiable component ($Var\left(\sum_{b \in B_s} w^s_{bt-1}(f_{bt} + \nu_{sbt})\right)$) and the variance of the non-diversifiable component ($Var\,(f_{st})$). "ln Herfindahl" is the Herfindahl of sales across buyers, computed the first year the firm appears in the data, while "ln Herf. ac. prod." is the Herfindahl across products. "ln value of exports" is the (initial) trade value. "ln # years" is the number of periods the firm is observed (which varies between 4 and 12). "Entrant" and "Young exporter" are dummy variables equal to one if the firm just entered the market when observed for the first time, or entered it less than two years before. The coefficients are identified in relative terms with respect to mature exporters. "$\mathbb{1} = 1$ if HQ in dest." and "$\mathbb{1} = 1$ if aff. in dest." proxy the extent of intra-firm trade flows.

Table 10: Summary statistics on the variance components, at the aggregate level

	(1) Mean	(2) Std.Dev	(3) Contrib.	(4) Partial Corr.
Variance of				
Aggregate growth g_t	0.0056	0.0033		
Macro component f_{Ct}	0.0003	0.0000	0.064	0.041^{***}
Seller-specific component f_{st}	0.0020	0.0013	0.303	0.301^{***}
Buyer-specific component f_{bt}	0.0028	0.0014	0.464	0.458^{***}
Match-specific residual ν_{sbt}	0.0039	0.0036	0.445	0.718^{***}
Buyer-specific input cost $BSIC_{bt}$	0.0008	0.0007	0.113	0.154^{***}
Covariance terms Cov	-0.0042	0.0043	-0.660	-0.672^{***}
Count observations	11			

Notes: This table gives summary statistics on the variance of aggregate growth to one destination and its components. The decomposition is based on equation (15), where the "Cov" term is the sum of all covariance terms involving f_{Ct}, f_{st}, f_{bt}, ν_{sbt} and $BSIC_{bt}$, both within and across seller-buyer pairs. Columns (1) and (2) respectively report the mean and standard deviation of the overall variance, computed over destination countries. Column (3) is the median contribution of each variance component to the total variance (eg. $Med(var(f_{Ct})/var(g_t))$). Column (4) is the partial correlation between each variance component and the overall variance. *** indicates significance at the 1% level.

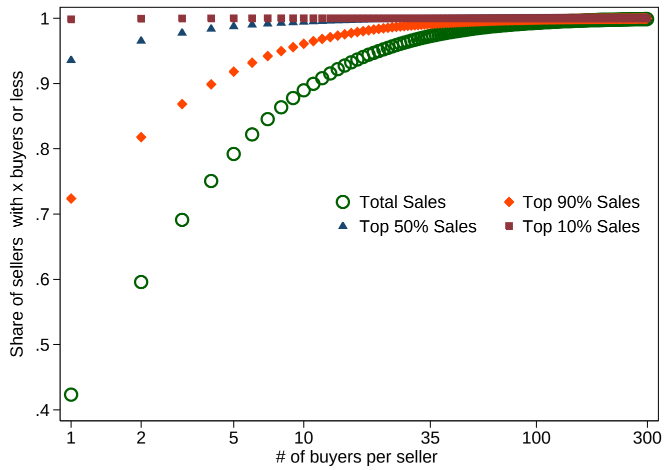
Figure 1: Cumulated share of exporters, importers and seller-buyer pairs in the total value of EU11 exports



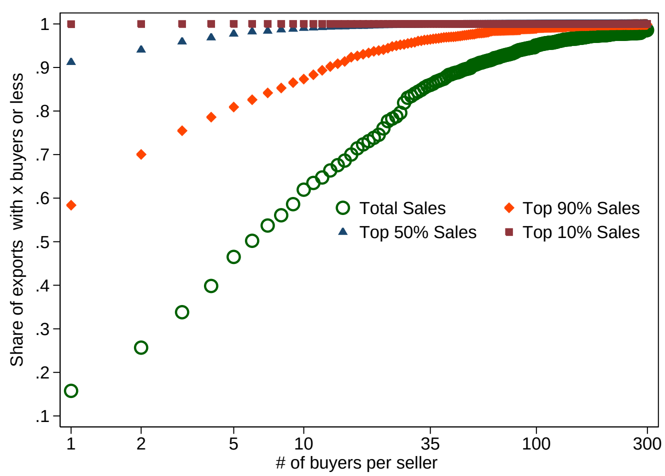
Notes: Each graph displays the share of the total value of French exports to EU11 countries that is attributable to the **x**% smallest individuals, in 2007. The first panel cumulates the sales of exporters of increasing size. The second panel cumulates importers' purchases. The third panel cumulates exporter-importer transactions. For instance, the number that corresponds to the point 80 of the x-axis in the first graph reads as follows: The cumulated contribution of the 80% smallest exporters is around 7%.

Figure 2: Number of Buyers per Seller

Share of sellers

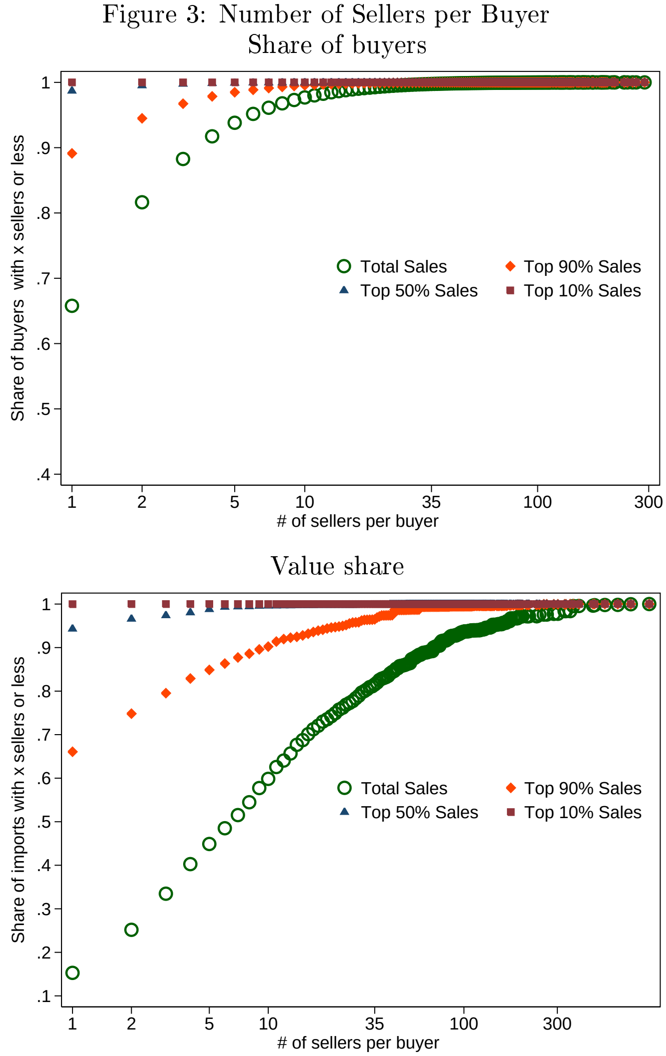


Value share



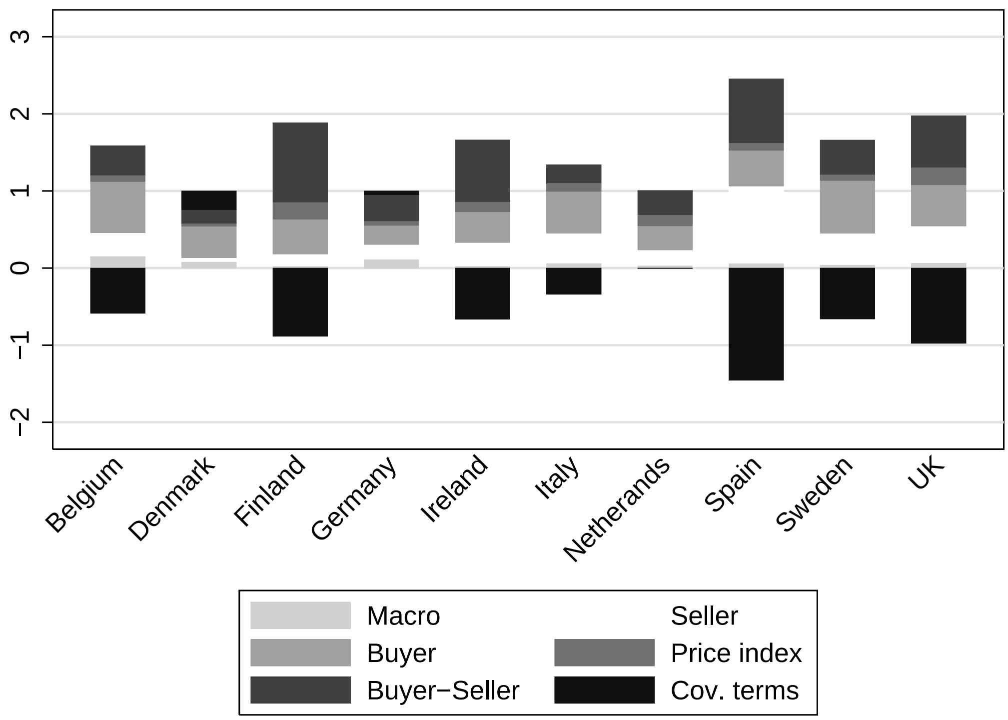
Notes: Proportion of sellers (top panel) and share of trade accounted for by sellers (bottom panel) that serve x buyers or less in a given destination, in 2007. The green circles correspond to total exports. The distributions labeled "Top X% Sales" are computed restricting the amount of each firm's sales to the X first percentiles of the distribution of sales when transactions are ordered by the decreasing share of the buyer in the firm's total sales. The line in red for instance interprets as follows: If, for each exporter, we neglect the set of the smallest buyers contributing to the last 10% of the exporter's market-specific sales, more than 70% of exporters have a degree of one buyer while only 5% have 10 buyers or more.

Figure 3: Number of Sellers per Buyer



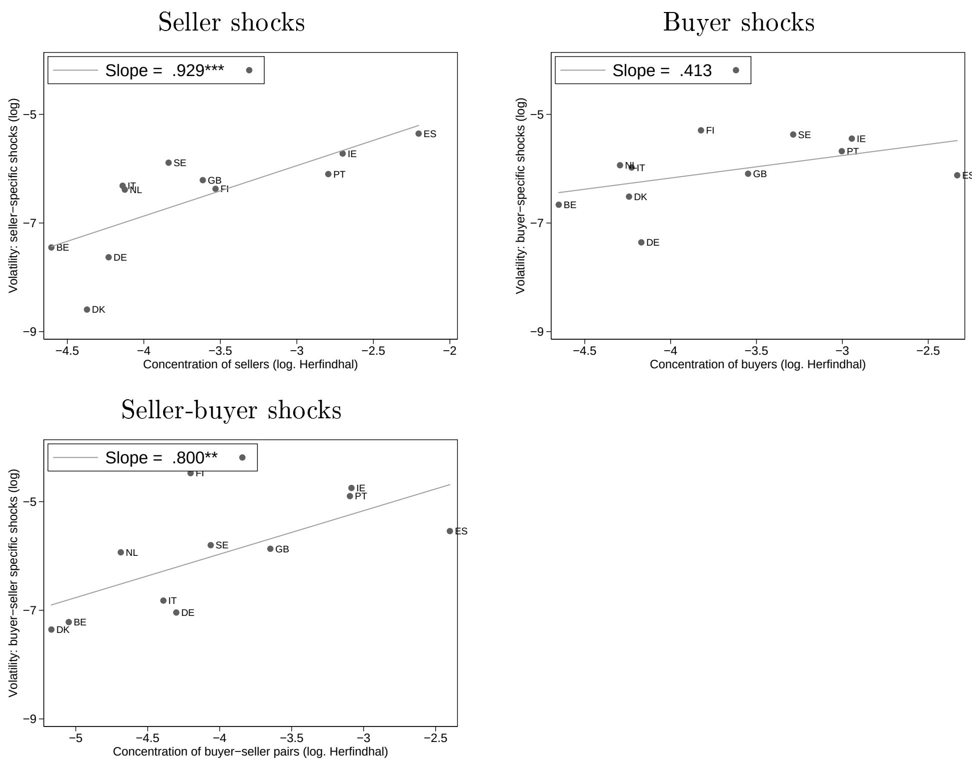
Notes: Proportion of buyers (top panel) and share of trade accounted for by buyers (bottom panel) that purchase from x sellers or less. Statistics based on all EU11 bilateral export flows in 2007. The green circles correspond to total imports. The distributions labeled "Top X% Sales" are computed restricting the amount of each firm's purchases to the X first percentiles of the distribution of purchases when transactions are ordered by the decreasing share of the seller in the firm's total purchases.

Figure 4: Contribution of the shocks to the variance of sales, by destination



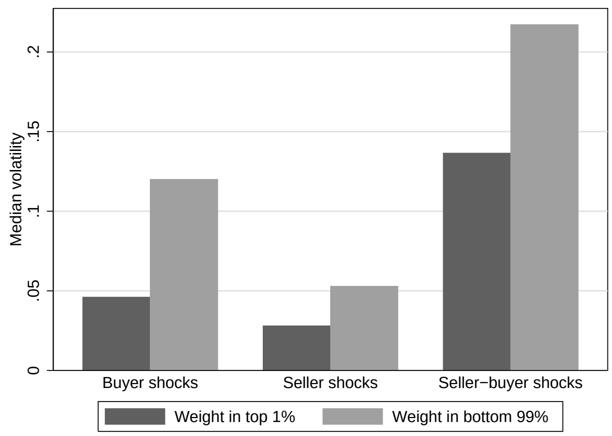
Notes: The graph displays the six components of the variance in the large, for each destination country. Portugal is neglected because of its status of outlier in this graph: Its overall variance is close to zero (0.002) which implies that the negative covariance of -.01 represents a huge relative contribution.

Figure 5: Concentration of trade flows and granular components



Notes: The graphs plot each of the components of the variance in the large induced by one type of individual shocks against the concentration of sales in the corresponding dimension. The slope of the regression line is reported in the legend with *** indicating significance at the 1% level.

Figure 6: Volatility of small and large firms and transactions



Notes: The graph plots the seller, buyer and seller-buyer components of the volatility for small and large firms. Firm/transaction size is measured by the average weight in bilateral French trade. The top 1% consists of the firms/transactions whose weight belongs to the first percentile in terms of the weight in bilateral exports. The bottom 99% is the rest. Within each group, the volatility of is defined as the median variance of each shock.

Table A.1: Coverage

	Value of exports (billion euros)	# of observations
All	2,180	14,069,787
Enough obs to compute g_{st}	1,960	12,093,470
Intensive margin	1,800	7,209,663
Excluding outliers ($g_{st} \in [-0.8; 4]$)	1,670	6,025,288
All shocks identified under restriction (10)	1,560	5,811,303
Enough obs to compute $Var(g_{st})$	892	3,085,338

Notes: This table gives the coverage of the sample used in the empirical analysis depending on the restrictions we apply.

Table A.2: Summary statistics on the intensive and extensive margins of firm-level volatility

	(1) Mean	(2) Std. Dev	(3) Contrib.	(4) Partial Corr.
Variance of				
Firm growth $\tilde{g}_{st}$	0.4330	0.6829		
Intensive component g_{st}	0.3257	0.4574	0.920	0.523***
Extensive component $g_{st}^{Ext.}$	0.2010	0.5161	0.263	0.514***
Covariance term $Cov(g_{st}, g_{st}^{Ext})$	-0.0937	0.4344	-0.103	-0.038***
Count observations	52,831			

Notes: This table gives summary statistics on the variance of firm growth and its extensive and intensive components. The decomposition is based on equation (A.3), where the "Cov" term is the covariance between g_{st} and g_{st}^{Ext}. Column (1) reports the mean variance in the population of firms, Column (2) its standard deviation. Column (3) is the median contribution of each variance component to the total variance (eg. $Med(var(g_{st})/var(g_{st}^{Tot})))$. Column (4) is the partial correlation between each variance component and the overall variance. The sample is restricted to variances computed on at least four growth rates. *** indicates significance at the 1% level.

References

Abowd, John, Francis Kramarz, Paul Lengermann, and Perez-Duarte Sebastien, "Are Good Workers Employed by Good Firms? A Simple Test of Positive Assortative Matching Models for France and the United States," mimeo, CREST 2004.

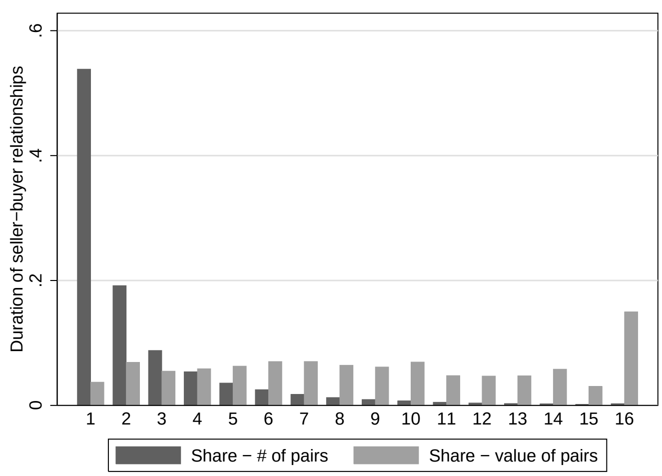
Abowd, John M., Francis Kramarz, and David N. Margolis, "High Wage Workers and High Wage Firms," *Econometrica*, March 1999, *67* (2), 251–334.

Table A.3: Summary statistics on the margins, at the aggregate level

	(1) Mean	(2) Std. Dev	(3) Contr.	(4) Partial Corr.
Variance of				
Growth $\tilde{g}_t$	0.0066	0.0076		
Intensive component g_t	0.0065	0.0069	1.073	0.928***
Extensive component seller $g_t^{Ext-seller}$	0.0010	0.0012	0.142	0.155***
Extensive component buyer $g_t^{Ext-buyer}$	0.0005	0.0006	0.128	0.050***
Covariance term $Cov.$	-0.0015	0.0024	- 0.287	-0.134***
Count observations	11			

Notes: This table gives summary statistics on the variance of firm growth and its extensive and intensive components. The decomposition is based on equation (A.4), where the "Cov" term is covariance term involving g_t, $g_t^{Ext-seller}$ and $g_t^{Ext-buyer}$. Column (1) reports the mean variance in the population of firms, Column (2) its standard deviation. Column (3) is the median contribution of each variance component to the total variance (eg. $Med(var(g_t)/var(g_t^{Tot}))$). Column (4) is the partial correlation between each variance component and the overall variance. The sample is restricted to variances computed on at least four growth rates. *** indicates significance at the 1% level.

Figure A.1: Duration of seller-buyer relationship



Notes: The graph displays the distribution of seller-buyer durations (in years) in the dataset, both in proportion to the total number of pairs (dark grey bars) and in proportion to the total value of exports (light grey bars).

—, Robert H. Creecy, and Francis Kramarz, "Computing Person and Firm Effects Using Linked Longitudinal Employer-Employee Data," Longitudinal Employer-Household Dynamics Technical Papers 2002-06, Center for Economic Studies, U.S. Census Bureau March 2002.

Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi, "The Network Origins of Aggregate Fluctuations," *Econometrica*, September 2012, *80* (5), 1977–2016.

Andrews, M. J., L. Gill, T. Schank, and R. Upward, "High wage workers and low wage firms: negative assortative matching or limited mobility bias?," *Journal of the Royal Statistical Society Series A*, 2008, *171* (3), 673–697.

Arkolakis, Costas, "A Unified Theory of Firm Selection and Growth," NBER Working Papers 17553, National Bureau of Economic Research, Inc October 2011.

Asker, John, Allan Collard-Wexler, and Jan De Loecker, "Dynamic Inputs and Resource (Mis)Allocation," *Journal of Political Economy*, 2014, *122* (5), pp. 1013–1063.

Berman, Nicolas, Philippe Martin, and Thierry Mayer, "How do Different Exporters React to Exchange Rate Changes?," *The Quarterly Journal of Economics*, 2012, *127* (1), 437–492.

Bernard, Andrew, Andreas Moxnes, and Karen Helene Ulltveit-Moe, "Two-sided Heterogeneity and Trade," 2014.

Bernard, Andrew B., J. Bradford Jensen, Stephen J. Redding, and Peter K. Schott, "Wholesalers and Retailers in US Trade," *American Economic Review*, 2010, *100* (2), 408–413.

—, Stephen J. Redding, and Peter K. Schott, "Multiproduct Firms and Trade Liberalization," *The Quarterly Journal of Economics*, 2011, *126* (3), 1271–1318.

Bilbiie, Florin O., Fabio Ghironi, and Marc J. Melitz, "Endogenous Entry, Product Variety, and Business Cycles," *Journal of Political Economy*, 2012, *120* (2), 304 – 345.

Blaum, Joaquin, Claire Lelarge, and Michael Peters, "The Intensive Margin of Imports and Firm Productivity," 2013 Meeting Papers 525, Society for Economic Dynamics 2013.

Bricongne, Jean-Charles, Lionel Fontagné, Guillaume Gaulier, Daria Taglioni, and Vincent Vicard, "Firms and the global crisis: French exports in the turmoil," *Journal of International Economics*, 2012, *87* (1), 134–146.

Campbell, John Y., "Have Individual Stocks Become More Volatile? An Empirical Exploration of Idiosyncratic Risk," *Journal of Finance*, 02 2001, *56* (1), 1–43.

Carballo, Jerónimo, Gianmarco Ottaviano, and Christian Volpe Martincus, "The Buyer Margins of Firms' Exports," CEPR Discussion Papers 9584, C.E.P.R. Discussion Papers August 2013.

Caselli, F., M. Koren, M. Lisichy, and S. Tenreyro, "Diversification through Trade," 2015.

Chaney, Thomas, "The Network Structure of International Trade," *American Economic Review*, 2014, *104* (11), 3600–34.

Comin, Diego A. and Thomas Philippon, "The Rise in Firm-Level Volatility: Causes and Consequences," in "NBER Macroeconomics Annual 2005, Volume 20" NBER Chapters, National Bureau of Economic Research, Inc, March 2006, pp. 167–228.

Comin, Diego and Sunil Mulani, "Diverging Trends in Aggregate and Firm Volatility," *The Review of Economics and Statistics*, 2006, *88* (2), 374–383.

Davis, Steven J., John Haltiwanger, Ronald S. Jarmin, C.J. Krizan, Javier Miranda, Alfred Nucci, and Kristin Sandusky, "Measuring the Dynamics of Young and Small Businesses: Integrating the Employer and Nonemployer Universes," in "Producer Dynamics: New Evidence from Micro Data" NBER Chapters, National Bureau of Economic Research, Inc, March 2009, pp. 329–366.

Decker, Ryan, John Haltiwanger, Ron Jarmin, and Javier Miranda, "The Role of Entrepreneurship in US Job Creation and Economic Dynamism," *Journal of Economic Perspectives*, 2014, *28* (3), 3–24.

di Giovanni, Julian, Andrei A. Levchenko, and Isabelle Mejean, "Firms, Destinations, and Aggregate Fluctuations," *Econometrica*, 07 2014, *82* (4), 1303–1340.

Eaton, Jonathan, Marcela Eslava, C. J. Krizan, Maurice Kugler, and James Tybout, "A Search and Learning Model of Export Dynamics," 2013.

_ , Samuel Kortum, and Francis Kramarz, "An Anatomy of International Trade: Evidence From French Firms," *Econometrica*, 09 2011, *79* (5), 1453–1498.

Feenstra, Robert C, "New Product Varieties and the Measurement of International Prices," *American Economic Review*, March 1994, *84* (1), 157–77.

Fort, Teresa C, John Haltiwanger, Ron S Jarmin, and Javier Miranda, "How Firms Respond to Business Cycles: The Role of Firm Age and Firm Size," *IMF Economic Review*, 2013, *61* (3), 520–559.

Foster, Lucia, John C. Haltiwanger, and Chad Syverson, "The Slow Growth of New Plants: Learning about Demand?," NBER Working Papers 17853, National Bureau of Economic Research, Inc February 2012.

_ , John Haltiwanger, and Chad Syverson, "Reallocation, Firm Turnover, and Efficiency: Selection on Productivity or Profitability?," *American Economic Review*, March 2008, *98* (1), 394–425.

Gabaix, Xavier, "The Granular Origins of Aggregate Fluctuations," *Econometrica*, 05 2011, *79* (3), 733–772.

Hopenhayn, Hugo A, "Entry, Exit, and Firm Dynamics in Long Run Equilibrium," *Econometrica*, 1992, *60* (5), 1127–50.

J.M., Abowd, Crepon B., and Kramarz F., "Moment Estimation With Attrition: An Application to Economic Models," *Journal of the American Statistical Association*, December 2001, *96*, 1223–1231.

Kelly, Bryan, Hanno Lustig, and Stijn Van Nieuwerburgh, "Firm Volatility in Granular Networks," NBER Working Papers 19466, National Bureau of Economic Research, Inc September 2013.

Kleinert, Jörn, Julien Martin, and Farid Toubal, "The Few Leading The Many: Foreign Affiliates and Business Cycle Comovement," CEPR Discussion Papers 9129, C.E.P.R. Discussion Papers September 2012.

Koren, Miklós and Silvana Tenreyro, "Volatility and Development," *The Quarterly Journal of Economics*, 02 2007, *122* (1), 243–287.

Lee, Yoonsoo and Toshihiko Mukoyama, "Entry and exit of manufacturing plants over the business cycle," *European Economic Review*, 2015, *77* (0), 20 – 27.

Mayer, Thierry, Marc J. Melitz, and Gianmarco I. P. Ottaviano, "Market Size, Competition, and the Product Mix of Exporters," *American Economic Review*, February 2014, *104* (2), 495–536.

Oberfield, Ezra, "Business networks, production chains and productivity: A theory of input-output architecture," Working Paper Series WP-2011-12, Federal Reserve Bank of Chicago 2011.

Tenreyro, Silvana, Miklos Koren, and Francesco Caselli, "Diversification through Trade," 2012 Meeting Papers 539, Society for Economic Dynamics 2012.

Thesmar, David and Mathias Thoenig, "Contrasting Trends in Firm Volatility," *American Economic Journal: Macroeconomics*, 2011, *3* (4), 143–80.