Innovation and Product Reallocation in the Great Recession*

David Argente† University of Chicago

Munseob Lee‡ University of Chicago

Sara Moreira§ Northwestern University

February 2017

[Link to the latest version]

Very Preliminary and Incomplete

Abstract

We exploit detailed product- and firm-level data to study the sources of innovation and the patterns of productivity growth in the consumer goods sector over the period 2006-2014. Using a dataset that contains information on the product portfolio of each firm and the characteristics of each product, we document new facts on product reallocation. First, we find that reallocation of products happens within the boundaries of the firm, but the largest changes in product quality come from new firms launching new varieties and from small firms expanding to other product lines. Second, we document that product reallocation within and between firms are procyclical and that product reallocation within firms was more affected than product reallocation between firms during the great recession. Finally, we quantify the extent to which product reallocation affects firm-level productivity growth and and innovation as reflected by changes in their total factor productivity. Our preliminary findings suggest that within-firm product reallocation contributes substantially to the growth in innovation across all sectors.

JEL Classification Numbers: E3, O3, O4.

Key words:

---

*This paper is currently in preparation for the Carnegie-Rochester-NYU Conference Series on Public Policy.

†Email: dargente@uchicago.edu. Address: 1126 E. 59th Street, Chicago, IL 60637.

‡Email: munseob@uchicago.edu. Address: 1126 E. 59th Street, Chicago, IL 60637.

§Email: sara.moreira@kellogg.northwestern.edu. Address: 2001 Sheridan Rd, Evanston, IL 60208.

1 Introduction

For decades, economists have identified product entry and exit as the key mechanism through which product innovation translates into economic growth (Aghion et al. (2014)). But despite the important theoretical implications of product innovation, little is known empirically about the processes of creation and destruction of products, and how these processes differ across different types of firms. In this paper we study product reallocation across and within producers and how it evolved during the Great Recession. What is the role of product reallocation to output growth and quality improvements in the recent decade? How sensitive is innovation by new firms, small incumbents, and large incumbents to changes in aggregate economic conditions? New evidence on these questions will help understand how resources are allocated to their best use within an economy and inform the recent debate on the sources of productivity slow down in the U.S. (Davis and Haltiwanger (2014), Decker et al. (2014)).

In this paper, we exploit detailed product- and firm-level data to study the sources of innovation and the patterns of productivity growth in the consumer goods sector over the period 2006Q3–2014Q2. Using a dataset that contains information on the products of each firm and the characteristics of each product, we document new facts on product reallocation. First, we find that an important component of reallocation of products happens within the boundaries of the firm. Second, the largest changes in product quality come from new firms launching new varieties and from small firms expanding to other product lines. Third, we document that product reallocation within firms is procyclical. Fourth, we find that within-firm product reallocation is larger in high productive firms and firms that invest more in R&D. Finally, we quantify how important how product reallocation affects firm-level productivity growth and and innovation as reflected by changes in their total factor productivity.

We rely primarily on the Nielsen Retail Measurement Services scanner dataset, which consists of more than 100 billion observations of weekly prices, quantities, and store information generated by point-of-sale systems. The data consist of approximately 1.4 million distinct products identified by a 12-digit universal product code (UPC). We combine information of prices with the weight and volume of the product to compute unit values in order to approximate the quality of each product.[1] In addition, we identify the firm owning each UPC by obtaining information from GS1, the single official source of barcodes in the United States. The GS1 data contains the list of all U.S. barcodes along with information on the firms that own them. Our combined dataset allows us to have revenue, price, quantity, and quality for each product of a firm's portfolio and allow us to study how the within and between margin of product creation and destruction evolve over time.

[1]The data is organized into 1,235 product categories that range from food to computer software and cover around 40% of all of the expenditures on goods in the CPI.

We start our empirical analysis by examining the amount of innovation produced by three different types of firms: i) existing firms improving on their own products, ii) existing firms improving on products by other firms, and iii) new firms entering the market with new brands and varieties. Consistent with the previous literature we find that most product reallocation occurs within the boundaries of the firm. We also find, however, that the largest quality improvements on existing products come from firms launching products in different product categories or from new firms introducing new varieties in the market. In fact, new firms launch fewer but more innovative products.

In order to understand the source of these innovation patterns, we proceed to analyze the patterns of product reallocation during the Great Recession for our three categories of firms.[2] Unlike previous studies, we document that during the great recession both product creation and destruction show procyclical patterns. As a result, the degree of product reallocation decreased substantially during the Great Recession. This is especially true for within-firm product reallocation.

Finally, to understand and quantify the relevance of product reallocation on innovation activity, we perform a simple accounting decomposition of quality dynamics in each product category. Our decomposition of the change in product quality in a given category allows us to consider the roles of the changing shares of economic activity of multi-product firms and new single-product firms. Our findings suggest that reallocation effects account for a large share of the changes in product quality growth over this period. While between-firms reallocation is significant, the majority of the changes in quality growth can be explained by within-firm reallocation. Our findings suggest that industrial and innovation policies that aim at increasing economic growth should contemplate the relative importance of the product-mix decisions of both incumbents and new firms.

Despite the vast theoretical implications of product reallocation, the empirical analysis on the aggregate behavior of product turnover lags far behind its theoretical counterpart due to data limitations. The extant literature on reallocation has focused on the input markets using establishment and labor market data (e.g. Davis and Haltiwanger (1992), Foster et al. (2001) and Foster et al. (2006)). By contrast, we focus on reallocation in output markets. Our work also relates to Bernard et al. (2010), who study the extent of product switching within firms using production classification codes (five-digit SIC codes). We substantially improve on this paper by measuring products at a much finer level using scanner data. Broda and

[2] The focus on multi-product firms' product mix decisions is relevant to the extent that the changes in the product mix account for a significant portion of changes in firms' output over time. We document that product churning/reallocation accounts for a sizable share of the increase in output in customer goods industries, which suggests that product mix changes represent a potentially important channel through which resources move from less to more efficient uses.

Weinstein (2010) studies the patterns of product entry and exit using a similar dataset to ours, which is collected from consumers rather than stores. Collecting data at the store level offers the advantage of observing, for the categories available, the entire universe of products for which a transaction is recorded in a given week rather than the products consumed by a sample of households. Therefore, our dataset allows us to cover less frequently consumed goods and provides less noisy measures of entry and exit of a product. Our work also differs substantially from Broda and Weinstein (2010) in that we subdivide firms into different categories and study how their characteristics affect their product reallocation patterns. For instance, we examine multi-product and single-products firms separately to analyze within-firm and between-firm reallocation separately. We also distinguish firms according to their level of diversification (i.e. firms selling in a single product category versus firms selling in multiple product categories). Finally, we focus particularly on the business cycle dynamics of product reallocation. To the best of our knowledge, our study is the first to study product creation and destruction during the Great Recession. Lastly, by studying the connection between reallocation and quality growth, our work links studies on reallocation, which focus mainly on moving resources from less to more efficient uses to enhance productivity growth, to the parallel literature on innovation (e.g. Klette and Kortum (2004), Acemoglu et al. (2013), Akcigit and Kerr (2010), Garcia-Macia et al. (2016)).

2 Data Description

2.1 Baseline Product-Level Dataset

We rely primarily on the Nielsen Retail Measurement Services (RMS) scanner dataset, which is provided by the Kilts-Nielsen Data Center at the University of Chicago Booth School of Business. The RMS consists of more than 100 billion unique observations at the week $\times$ store $\times$ UPC level.[3] Each individual store reports weekly prices and quantities sold for every UPC code that had any sales volume during the week.

The data is generated by point-of-sale systems and contains approximately 40,000 distinct stores from 90 retail chains across 371 MSAs and 2,500 counties between January 2006 and December 2014. The data set includes around 12 billion transactions per year worth, on average, \$220 billion. Over our sample period the total sales across all retail establishments are worth approximately \$2 trillion and represent 53% of all sales in grocery stores, 55% in drug stores, 32% in mass merchandisers, 2% in convenience stores, and 1% in liquor stores.

[3] A Universal Product Code (UPC) is a code consisting of 12 numerical digits that is uniquely assigned to each specific item.

The baseline data consist of approximately 1.64 million distinct products identified by a unique UPC. The data is organized into 1,070 detailed product categories, 114 aggregated product categories, and 10 major departments.[4] Each UPC contains information on the brand, size, packaging, and a rich set of product features including the weight and the volume of the product which we use to compute unit values. Our data set combines all sales at the national and quarterly level. For each product j in quarter t, we define revenue r_{jt} as total revenue across all stores and weeks of the quarter. Likewise, quantity q_{jt} is defined as total quantities sold across all stores and weeks of the quarter. Price p_{jt} is defined by the ratio of revenue to quantity, which is equivalent to quantity weighted average price.

The identification of entries and exits is critical to our analysis. For each product we use the panel structure to measure the entry and exit periods. We follow Broda and Weinstein (2010) and Argente and Yeh (2016) and use the UPC as the main product identifier. This is because it is rare that a meaningful quality change occurs without resulting in a UPC change. Two concerns may arise from this assumption. Chevalier et al. (2003) notes that some UPCs might get discontinued only to have the same product appear with a new UPC. This is not a concern in our data set because Nielsen detects these UPCs and assigns them their prior UPC. The second concern is that companies could recycle UPCs and use them for different products. Nielsen also identifies these UPCs and assigns them a different UPC version, which we use to identify products.[5]

We define entry as the first quarter of sales of a product and exit as the first quarter after we last observe a product being sold.[6] To study the entry- and exit rates patterns we use information for all products in the period 2006Q3–2014Q2, that include cohorts born from 2006Q2 to 2014Q2 and cohorts born before that period, from whom we cannot determine the cohort and age. In addition, given that our estimates of product entry and exit may be affected by the entry and exit of stores in the sample, we consider only a balanced sample

[4]The ten major departments are: health and beauty aids (e.g. cosmetics, pain remedies), dry grocery (e.g. baby food, canned vegetables), frozen foods, dairy (e.g. cheese, yogurt), deli, packaged meat, fresh produce, non-food grocery (e.g. detergent, pet care), alcohol and general merchandise (e.g. batteries/flashlights, computer/electronic).

[5]For robustness, we construct broader definitions of a good using the product attributes included in our data set as in Kaplan and Menzio (2014). Specifically, we defined goods as the sets of products that share the same features, the same size and the same brand, but may have different barcodes. All the findings described below are robust to this alternative definition.

[6]A product is first observed in the RMS dataset when a store sells it for the first time. Firms often do not introduce goods in different stores at the same time, and thus, when we first observe a new UPC it may take some time until it is disseminated across stores. Likewise, firms may stop producing a good in a certain time and we can still see some transactions after that, as stores may still have some inventories. In the Appendix, we consider alternative definitions, where entry is defined as the second quarter of sales of a good, and exit as the second quarter to the last that we observe a product being sold. Relative to the definition of entry (exit) as the first (last) quarter where sales are registered, this last measure is more likely to represent a full quarter of sales, instead of just a few weeks.

of stores during our sample period. Our final sample consists of 33,056 stores and a total an average of 425,000 different UPCs every year.

2.2 Matching Firm and Products

We link companies and products using information obtained from GS1 US, the single official source of barcodes. In order to obtain a barcode, companies must first obtain a GS1 Company Prefix. The prefix is a 6-10 digits number that uniquely identifies companies and their products in over 100 countries where GS1 is present.[7] The GS1 US data contains the universe Company Prefixes generated in the US along with information of the firms including their location. We combine these prefixes with those present in each UPC code in the RMS.[8] Our combined data set allows us to compute revenue, price, quantity, and quality for each product in a firm's portfolio and allow us to study how the within and between margins of product creation and destruction evolve over time. Appendix A.1 provides more information on how we use the product level dataset and the information on the firm, to construct our dataset.

3 Measurement of Reallocation

3.1 Aggregated Reallocation

Using information on the number of new products, exit products, and total products by each firm i over time t, we define the aggregate entry rate, exit rate and reallocation as follows

$$n_t = \frac{\sum_i N_{it}}{\sum_i T_{it}} \tag{1}$$

$$x_t = \frac{\sum_i X_{it}}{\sum_i T_{it-1}} \tag{2}$$

$$r_t = n_t + x_t \tag{3}$$

where N_{it}, X_{it}, and T_{it} are total number of entrant products, exit products and total products, respectively. This measure of aggregated reallocation captures the importance of entry and exit of products in the Consumer Goods Industries. Figure 1 shows that on average every quarter there are 7.5 percent of products that are being reallocated in the economy, with

[7]The number of digits in a Company Prefix indicates different capacities for companies to create barcodes. A 10 digits Company Prefix allows firms to create 10 unique barcodes and a 6 digits prefix allows them to create up to 100,000 unique items. These numbers, combined with a check digit, form the UPC code.

[8]Less than 5 percent of the UPCs belong to prefixes not generated in the US. We were not able to find a firm identifier for these products.

almost equal contribution for the entry and exit of products. Over the period under analysis, the quarterly reallocation is very procyclical. It declined from around 8 per cent in 2006, to 6.5 per cent in 2009, and back to 8 per cent in 2013. Most of the this pattern in reallocation results from the strong procyclicality in the aggregated entry rate (declined by almost 50 per cent in the period 2006 to 2009), while the exit rate is increasing over the entire entire period. .

To better understand the sources of reallocation, we decompose the aggregate entry and exit rate by types of firms. We classify firms by their net creation of products (expanding, contracting and stable), and access to the market (entrant, exit, incumbent).[9] Figures X show that 80 percent of the aggregated product creation is made by incumbent firms that are expanding the number of products. Entrant firms, only add 6 percent of the total products created, and the remaining comes from firms that create products but that are contracting. Over the period 2006Q3 and 2014Q2, the strong decline in the entry rate during the great recession is pervasive across all type of firms. Figure X shows that product destruction is more common among firms that are contracting (around 65 percent). However, expansionary firms are still important in the destruction of products, since almost 25 percent of products are exiting the market are from firms that are expanding. Overall, most of the reallocation of products comes from incumbent firms that are expanding (Figure X), and the slow down in the reallocation during the great Recession is common across the different types of firms.

There is substantial heterogeneity in the size of firms that produce consumer goods products. We classify firms into their quartile of revenue and we measure the contribution of each to the overall aggregate reallocation. Figure X shows that most reallocation happens among the top quartile of the distribution of revenue. Over the recent recession, all four groups of firms show a substantial decline in the reallocation from 2006 to 2010, followed by a recovery until 2014. The declines in the reallocation are larger than 50 percent among smaller firms, and around 15 percent for the top quartile.

Another important source of heterogeneity in this industry is the degree of diversification of products between firms (see details in Appendix) Diversified firms have a higher contribution in the reallocation of products (Figure X). In the recent recession both diversified and non-diversified firms contributed to the procyclicality of the aggregated reallocation.

3.2 Within Firm Reallocation

Using information on the number of new products, exit products, and total products by each firm i over time t, we define the average within firm entry rate, exit rate and reallocation as

[9] Appendix A.2 provides details on the disaggregation.

follows

$$\bar{n}_t = \frac{1}{firms_t} \sum_i n_{it} \tag{4}$$

$$\bar{x}_t = \frac{1}{firms_{t-1}} \sum_i x_{it} \tag{5}$$

$$\bar{r}_t = \bar{n}_t + \bar{x}_t \tag{6}$$

This measure of within firm reallocation captures average rate of reallocation of firms. Figure 2 shows that on average every quarter, firms reallocate 9.5 percent of their products, with almost equal contribution for the entry and exit of products. Over the period under analysis, the quarterly within-firms reallocation is very procyclical. It declined from around 11 per cent in 2006, to 8 per cent in 2009, and back to 11 per cent in 2013. As with, the aggregated reallocation, most of the decline in reallocation within firms resulted from the decline in the creation of products during the Great Recession.

t.b.c

3.3 Decomposition of aggregated reallocation

3.3.1 Olley and Payes (1996)

Aggregate entry rate as a function of firm's specific entry rates:

$$r_t = \sum_i \frac{R_{it}}{T_{it}} \frac{T_{it}}{\sum_i T_{it}} = \sum_i r_{it} t_{it} = \frac{1}{\Gamma_t} \sum_{i=1}^{\Gamma_t} r_{it} + \sum_{i=1}^{\Gamma_t} r_{it}(t_{it} - \frac{1}{\Gamma_t}) = \bar{n}_t + \sum_{i=1}^{\Gamma_t} (r_{it} - \bar{r}_t)(t_{it} - \bar{t}_t) \tag{7}$$

The first component is the cross-sectional (unweighted) meand across all firms, the second component measures where more products are located among firms with high or low entry rates.

The changes can be written as

$$r_t - r_0 = \bar{r}_{it} - \bar{r}_{i0} + \sum_{i=1}^{\Gamma_t} (r_{it} - \bar{r}_t)(t_{it} - \bar{t}_t) - \sum_{i=1}^{\Gamma_0} n_{i0}(r_{i0} - \bar{n}_0)(t_{i0} - \bar{t}_0) \tag{8}$$

The first component represents changes in the average entry rate within firm, and the second component is the adjustment by differences in size across firms.

tbc

4 Reallocation and Productivity

Does product reallocation matter for productivity growth? In most models of creative destruction, the pace of output reallocation plays an important role to determine productivity dynamics. These models emphasize that adopting new products inherently involves the destruction of the old ones. This process could occur through firm entry and exit, through the improvement of existing products, or through the required trial and error that developing a product often involves.

The relationship between product reallocation and TFP has been, nonetheless, hard to test empirically due to the lack of data combining both information on the firm's products and their productivity. In this section, we describe how we solve this issue by merging the Nielsen data with the Compustat database. In addition, we show that there exists a positive relationship between our measure of product reallocation and TFP and we present evidence that this relationship depends crucially on the innovation efforts of the firm.

4.1 Matching Nielsen RMS and Compustat

To combine the Nielsen data with the Compustat database, we matched the names provided by the GS1 to those in Compustat using the string matching algorithm described in Schoenle (2016). After applying the algorithm we matched 435 publicly traded companies over our sample period. Importantly, in some cases there are many firms in the RMS matched to a single corporation in Compustat. This is largely the result of past mergers and acquisitions. For example, Gillette has its own company prefix in GS1 because it possessed its own barcodes before it was acquired by Procter and Gamble. Because our unit of analysis will be the firms in the Compustat database, in what follows we will refer to each public corporation as firms and to each owner of a GS1 code as subsidiary. The median firm in Compustat has only one subsidiary in the RMS and the mean only 2.3.[10]

Our matched sample represents 26% of the total sales in Compustat and 39% of the total revenue in the RMS. Approximately 17% of the total number of products in the data belong to publicly traded firms. Table I presents summary statistics of both the Compustat data and our matched sample for a number of firm characteristics. Our sample mainly consists of large major U.S. companies with mean annual sales of $10.8 billion. These firms are on average more productive, have higher credit rating, and have less leverage. Nonetheless, the merged data set represents well publicly traded firms in many dimensions such as: market power (price to margin ratio), volatility of demand and costs, share of labor expenses, research and

[10] The version of the GS1 data we have accounts for all mergers and acquisitions that occurred before October 2013.

development expenses, and financial variables such as cash flows and dividends. We describe in detail the construction of this variables in Appendix A.3.

Likewise, our matched sample represents well the various measures of product creation we described in the previous sections. Table II shows that the entry rate, reallocation rate, and exit rate are extremely close in both data sets. Nonetheless, the firms in our sample sell on average more products and generate more revenue per year. To rule out the possibility that firm size or other type of heterogeneity drives our results, in the analysis that follows we control for an extended set of observables at the firm level.

4.2 Reallocation and Firm Productivity

We begin by documenting the relationship between our measure of product reallocation and firm level productivity. Table III and Table IV reports results from regressing the natural logarithm of TFP on the annual average of reallocation (measured on a quarterly basis). We use the following specification:

$$TFP_{i,f,t+1} = \alpha + \beta Reallocation_{i,f,t} + \Gamma X_{f,t} + \mu_f + \lambda_t + \epsilon_{i,f,t} \tag{9}$$

where i is the subsidiary, f is the firm, and t is the year. Our main focus is on *beta* which captures the direct impact of reallocation on firm productivity. $X_f t$ is a vector of firm level controls, μ_f firm (subsidiary) fixed effects and λ_t represents time effects. All our specifications include year-effects to control for possible trends and either firm or subsidiary fixed effects to control for other types of heterogeneity. Column (1) shows that both variables are strongly correlated even after controlling for the size of the firm. This is important because larger firms are more likely to have both higher rate of reallocation and higher productivity. Column (2) includes controls for market power by including the price-to-cost margin as control. Column (3) includes the standard deviation of sales to control for the possibility that firms with faster sales growth have higher rates of product entry and exit. Lastly, in column (4) we control for differences in financial constraints across firms by including the Kaplan-Zingales index. Our results are robust to the inclusion of these controls and the estimates of *beta* remain very similar across specifications. This is not surprising given that we include both firm and time effects in every specification which account for almost 85% of all possible variation.

Columns (5)-(8) show the same specifications including subsidiary fixed effects. Our conclusions remain very similar: an increase of one standard deviation in reallocation increases TFP between 0.4 and 0.6%. Table IV shows our results when, instead of equally weighting firms, we weight them by their revenue. The table shows that our estimates for *beta* are almost four times which is expected given the large right-tail on the distribution of revenue that we observe in the data. Importantly, this estimates are both economically and statistically significant which shows that reallocation has an important impact on firm-level productivity.

4.3 Reallocation and R&D Expenses

What determines the rate of reallocation within firms? The answer of this questions seems crucial to uncover the determinants of growth given the importance of reallocation on total factor productivity. In many endogenous growth models, the simple presence of creative destruction increases the incentives for firms to increase their innovation efforts. For example, in Aghion and Howitt (1992) firms get monopoly rents for their innovations until the next innovation arrives. In this case, the incentives for investing in innovation are substantial.

To test the prediction of this models, we explore the relationship between reallocation and the innovation effort of firms proxied by their R&D expenditures. We use this measure because R&D expenses, as defined by Compustat, encompass all planned search aimed at the discovery of new knowledge that could lead to new products or the improvement of the existing ones. Our specification is the following:

$$Reallocation_{i,f,t+1} = \alpha + \beta Research_{f,t} + \Gamma X_{f,t} + \mu_f + \lambda_t + \epsilon_{i,f,t} \tag{10}$$

where $Research_{f,t}$ represents the ratio of research and development expenses to total assets for firm f at time t.

Table V reports our results. Once again we include both firm (subsidiary) fixed effects and and year effects in all our specifications. Column (1) shows that R&D expenditures have a large positive impact on reallocation after controlling for firm size; larger firms tend to engage in more R&D activities. Next, we add a wide range of controls to disentangle the effect of R&D from potentially confounding firm-level factors. In column (2) we include the price cost margin and in column (3) a control for firm idiosyncratic volatility. Our results do not vary under these specifications or if measures of financial constraints are included (column (4)).[11] The rest of the columns (5)-(8) show that the point estimate does not vary significantly if we include subsidiary fixed effects instead of firm fixed effects.

Finally, Table VI shows that, just as in the previous section, our estimates are larger if we weight firms by revenue instead of equally. The point estimate is large and statistically significant; a hypothetical increase in R&D expenditures relative to sales of 5 percentage

[11] The magnitude of the point estimate for β and its significance is not sensitive to including other measures of financial constraints such as the firm's S&P long-term issuer rating.

points increases the reallocation rate by 1-1.5 percentage points ($100 \times 0.05 \times 0.2$). This is equivalent to an increase of close to 15% in the reallocation rate and 2% in total factor productivity. Overall, our estimates suggest that an increase in the innovation efforts by the firm is associated with a substantial creative destruction as measured by product reallocation; this in turn could have a substantial impact for firm's productivity and growth.

5 Conclusion

References

Acemoglu, Daron, Ufuk Akcigit, Nicholas Bloom, and William R Kerr, "Innovation, Reallocation and Growth," *NBER Working Paper No. 18993*, 2013.

Aghion, Philippe and Peter Howitt, "A Model of Growth through Creative Destruction," *Econometrica*, 1992, *60* (2), 323–351.

_ , Ufuk Akcigit, and Peter Howitt, "What Do We Learn from Schumpeterian Growth Theory?," in Philippe Aghion and Steven N Durlauf, eds., *Handbook of Economic Growth*, Vol. 2, Oxford: North Holland, 2014, chapter 1, pp. 515–563.

Akcigit, Ufuk and William R Kerr, "Growth Through Heterogeneous Innovations," *NBER Working Paper No. 16443*, 2010.

Argente, David and Chen Yeh, "Price-Setting under Uncertain Demand: A Menu Cost Model with Price Experimentation," 2016.

Bernard, Andrew B, Stephen J Redding, and Peter K Schott, "Multiple-Product Firms and Product Switching," *American Economic Review*, 2010, *100* (1), 70–97.

Broda, Christian and David E Weinstein, "Product Creation and Destruction: Evidence and Price Implications," *American Economic Review*, 2010, *100* (3), 691–723.

Chevalier, Judith A, Anil K Kashyap, and Peter E Rossi, "Why Don't Prices Rise During Periods of Peak Demand? Evidence from Scanner Data," *American Economic Review*, 2003, *93* (1), 15–37.

Davis, Steven J and John Haltiwanger, "Gross Job Creation, Gross Job Destruction, and Employment Reallocation," *The Quarterly Journal of Economics*, 1992, *107* (3), 819–863.

_ and _ , "Labor Market Fluidity and Economic Performance," *NBER Working Paper No. 20479*, 2014.

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

Foster, Lucia, John Haltiwanger, and Cornell J Krizan, "Aggregate Productivity Growth. Lessons from Microeconomic Evidence," in "New Developments in Productivity Analysis," University of Chicago Press, 2001, pp. 303–372.

—, —, and —, “Market Selection, Reallocation, and Restructuring in the U.S. Retail Trade Sector in the 1990s,” *The Review of Economics and Statistics*, 2006, *88* (4), 748–758.

Garcia-Macia, Daniel, Chang-Tai Hsieh, and Peter J Klenow, “How Destructive is Innovation?,” 2016.

Gorodnichenko, Yuriy and Michael Weber, “Are sticky prices costly? Evidence from the stock market,” *The American Economic Review*, 2016, *106* (1), 165–199.

İmrohoroğlu, Ayşe and Şelale Tüzel, “Firm-level productivity, risk, and return,” *Management Science*, 2014, *60* (8), 2073–2090.

Kaplan, Greg and Guido Menzio, “The Morphology of Price Dispersion,” *NBER Working Paper No. 19877*, 2014.

Kaplan, Steven N and Luigi Zingales, “Do investment-cash flow sensitivities provide useful measures of financing constraints?,” *The Quarterly Journal of Economics*, 1997, *112* (1), 169–215.

Klette, Tor Jakob and Samuel Kortum, “Innovating Firms and Aggregate Innovation,” *Journal of Political Economy*, 2004, *112* (5).

Olley, G. Steven and Ariel Pakes, “The Dynamics of Productivity in the Telecommunications Equipment Industry,” *Econometrica*, 1996, *64* (6), 1263–1297.

Schoenle, Raphael, “International Menu Costs and Price Dynamics,” *Review of International Economics*, 2016.

6 Tables and Figures

Table I: Summary Statistics - Publicly Traded Firms vs Matched Sample

The table reports summary statistics for all publicly traded firms and those matched with the Nielsen RMS from 2006-2014. The construction of these variables is detailed in Appendix A.3. R&D along with the rest of the financial variables are reported relative to net sales. We report the natural logarithm of TFP. Total sales are reported in millions. All statistics are weighted by total sales except the number of firms, employees and sales.

	Full Data	Sample
Num. Firms	6265	435
Employees	9.96	30.21
Sales	3405.2	10834.5
R&D	0.03	0.03
TFP	0.04	0.07
Price Cost Margin	0.70	0.67
Share of Labor	0.13	0.13
Std. Sales	0.15	0.13
KZ	2.34	1.43
Cash	0.11	0.11
Cash Flow	0.10	0.11
Dividends	0.02	0.03
Leverage	0.96	0.81
Market Value Equity	1.04	1.20
Q Ratio	0.71	0.84
S&P	2.65	3.23

Table II: Summary Statistics - Nielsen RMS vs Matched Sample

The table reports summary statistics for all firms in the Nielsen RMS and those matched with the sample of publicly traded firms in Compustat from 2006-2014. Average modules (groups or departments) indicates the average number of modules (groups or departments) in which the firm sells products. Revenue main module (groups or departments) indicates the share of the firm's total revenue that comes from the module (groups or departments) from which the highest share of its revenue comes from. Revenue is reported in millions. The construction of the rest of the variables is detailed in main text. All statistics are weighted by revenue except the number of firms, number of products and revenue.

Variable	Full Data	Sample
Num. Firms	39997	1039
Num. Products	15.96	104.3
Revenue	1.10	16.46
Reallocation	0.08	0.08
Entry Rate	0.07	0.07
Exit Rate	0.04	0.04
Avg. Departments	2.34	2.66
Avg. Groups	9.12	11.10
Avg. Modules	30.1	29.21
Rev. Main Group	0.78	0.68
Rev. Main Module	0.58	0.49
Rev. Main Dept.	0.92	0.91

Table III: Reallocation and Firm-Level Productivity

The table reports the coefficients of OLS regressions with equal weights. The dependent variable is the natural logarithm of total factor productivity at the firm-level at t. Reallocation at $t-1$ is defined as the product entry rate plus the product exit rate at the subsidiary level as defined in the main text. The construction of the control variables is described in Appendix A.3. Other controls include firm fixed-effects and year fixed-effects (column (1)-column (4)) and subsidiary fixed-effects and year fixed-effects (column (5)-column (8)). The sample used include the set of publicly traded firms that are found in the Nielsen RMS. Standard errors are presented in parentheses, and are clustered at firm-level in columns (1)-(4) and at the subsidiary level in columns (5)-(8). ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reallocation	0.050**	0.049**	0.049**	0.045**	0.065*	0.065*	0.062*	0.058*
	(0.022)	(0.022)	(0.022)	(0.021)	(0.037)	(0.036)	(0.034)	(0.034)
Size	0.274***	0.250***	0.258***	0.255***	0.269***	0.244***	0.249***	0.247***
	(0.067)	(0.062)	(0.053)	(0.054)	(0.037)	(0.035)	(0.033)	(0.033)
Price Cost Margin		-1.483***	-1.437***	-1.340***		-1.450***	-1.401***	-1.307***
		(0.373)	(0.365)	(0.363)		(0.248)	(0.250)	(0.251)
Std. Sale			-0.436*	-0.446**			-0.448***	-0.457***
			(0.225)	(0.225)			(0.121)	(0.121)
Kaplan-Zingales				-0.000				0.000
				(0.001)				(0.001)
Observations	4,773	4,773	4,765	4,729	4,773	4,773	4,765	4,729
R-squared	0.844	0.852	0.856	0.856	0.849	0.855	0.860	0.859
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Effects	Yes	Yes	Yes	Yes	No	No	No	No
Subsidiary Effects	No	No	No	No	Yes	Yes	Yes	Yes

Table IV: Reallocation and Firm-Level Productivity (Revenue Weights)

The table reports the coefficients of OLS regressions with revenue weights. The dependent variable is the natural logarithm of total factor productivity at the firm-level at t. Reallocation at $t-1$ is defined as the product entry rate plus the product exit rate at the subsidiary level as defined in the main text. The construction of the control variables is described in Appendix A.3. Other controls include firm fixed-effects and year fixed-effects (column (1)-column (4)) and subsidiary fixed-effects and year fixed-effects (column (5)-column (8)). Revenue is winsorize at the 1% level. The sample used include the set of publicly traded firms that are found in the Nielsen RMS. Standard errors are presented in parentheses, and are clustered at firm-level in columns (1)-(4) and at the subsidiary level in columns (5)-(8). ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Reallocation	0.233*	0.212*	0.207*	0.193*	0.322*	0.299*	0.291*	0.273*
	(0.137)	(0.114)	(0.112)	(0.112)	(0.195)	(0.161)	(0.158)	(0.161)
Size	0.190**	0.160**	0.166**	0.160**	0.185**	0.155**	0.160**	0.155**
	(0.076)	(0.067)	(0.072)	(0.072)	(0.073)	(0.064)	(0.069)	(0.070)
Price Cost Margin		-1.137**	-1.135**	-1.076**		-1.142**	-1.139**	-1.082**
		(0.469)	(0.488)	(0.494)		(0.460)	(0.472)	(0.474)
Std. Sale			-0.143	-0.139			-0.136	-0.132
			(0.281)	(0.279)			(0.266)	(0.265)
Kaplan-Zingales				0.000				0.000
				(0.001)				(0.000)
Observations	4,773	4,773	4,765	4,729	4,773	4,773	4,765	4,729
R-squared	0.822	0.829	0.829	0.829	0.824	0.830	0.831	0.831
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Effects	Yes	Yes	Yes	Yes	No	No	No	No
Subsidiary Effects	No	No	No	No	Yes	Yes	Yes	Yes

Table V: Reallocation and R&D Expenses

The table reports the coefficients of OLS regressions with equal weights. The dependent variable reallocation at t, defined as the product entry rate plus the product exit rate, at the subsidiary level as defined in the main text. The main independent variable is the ratio of R&D expenses to total sales at $t-1$. The construction of the rest of the control variables is described in Appendix A.3. Other controls include firm fixed-effects and year fixed-effects (column (1)-column (4)) and subsidiary fixed-effects and year fixed-effects (column (5)-column (8)). The sample used include the set of publicly traded firms that are found in the Nielsen RMS. Standard errors are presented in parentheses, and are clustered at firm-level in columns (1)-(4) and at the subsidiary level in columns (5)-(8). ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
R&D	0.013*	0.023*	0.070***	0.070***	0.014*	0.023	0.072***	0.072***
	(0.007)	(0.014)	(0.023)	(0.023)	(0.008)	(0.015)	(0.024)	(0.024)
Size	-0.004	-0.004	-0.006	-0.005	0.000	0.000	-0.001	0.000
	(0.007)	(0.007)	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)	(0.007)
Price Cost Margin		-0.001	-0.004***	-0.004***		-0.001	-0.004***	-0.004***
		(0.001)	(0.001)	(0.001)		(0.001)	(0.002)	(0.002)
Std. Sale			0.009	0.008			0.012	0.011
			(0.021)	(0.021)			(0.021)	(0.021)
Kaplan-Zingales				-0.000				-0.000
				(0.000)				(0.000)
Observations	4,017	4,017	3,985	3,979	4,017	4,017	3,985	3,979
R-squared	0.275	0.275	0.276	0.267	0.470	0.470	0.471	0.465
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Effects	Yes	Yes	Yes	Yes	No	No	No	No
Subsidiary Effects	No	No	No	No	Yes	Yes	Yes	Yes

Table VI: Reallocation and R&D Expenses (Revenue Weights)

The table reports the coefficients of OLS regressions with equal weights. The dependent variable reallocation at t, defined as the product entry rate plus the product exit rate, at the subsidiary level as defined in the main text. The main independent variable is the ratio of R&D expenses to total sales at $t-1$. The construction of the rest of the control variables is described in Appendix A.3. Other controls include firm fixed-effects and year fixed-effects (column (1)-column (4)) and subsidiary fixed-effects and year fixed-effects (column (5)-column (8)). Revenue is winsorize at the 1% level. The sample used include the set of publicly traded firms that are found in the Nielsen RMS. Standard errors are presented in parentheses, and are clustered at firm-level in columns (1)-(4) and at the subsidiary level in columns (5)-(8). ***, **, and *, represent statistical significance at 1%, 5%, and 10% levels, respectively.

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
R&D	0.228**	0.255**	0.330**	0.331**	0.226*	0.251*	0.325**	0.326**
	(0.115)	(0.129)	(0.142)	(0.142)	(0.124)	(0.138)	(0.152)	(0.152)
Size	-0.004	-0.005	-0.005	-0.005	-0.005	-0.006	-0.007	-0.007
	(0.006)	(0.006)	(0.006)	(0.006)	(0.007)	(0.007)	(0.007)	(0.007)
Price Cost Margin		-0.011	-0.016	-0.016		-0.010	-0.015	-0.015
		(0.014)	(0.015)	(0.015)		(0.014)	(0.015)	(0.015)
Std. Sale			-0.026	-0.026			-0.026	-0.026
			(0.020)	(0.020)			(0.020)	(0.020)
Kaplan-Zingales				0.000				0.000
				(0.000)				(0.000)
Observations	4,017	4,017	3,985	3,979	4,017	4,017	3,985	3,979
R-squared	0.518	0.518	0.518	0.518	0.661	0.662	0.662	0.662
Year Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm Effects	Yes	Yes	Yes	Yes	No	No	No	No
Subsidiary Effects	No	No	No	No	Yes	Yes	Yes	Yes

Figure 1: Aggregate reallocation: the role of entry and exit

This figure plots the aggregated entry rate of products, exit rate of products, and reallocation as defined in the Appendix .

Figure 2: Reallocation within firms: the role of entry and exit

This figure plots entrant rate of products, exit rate of products, reallocation within firms each quarter from 2006Q3 to 2014Q2.

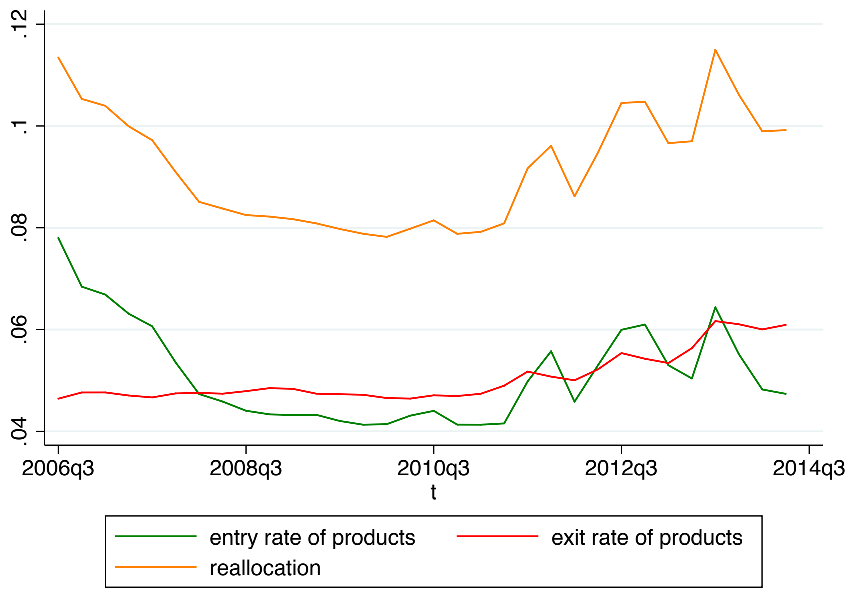


Figure 3: Total number of firms, total number of entrant firms, total number of exit firms. This figure plots total number of firms, total number of entrant firms, total number of exit firms each quarter from 2006Q3 to 2014Q2. The left axis is for total firms, and the right axis for entrants and exits.

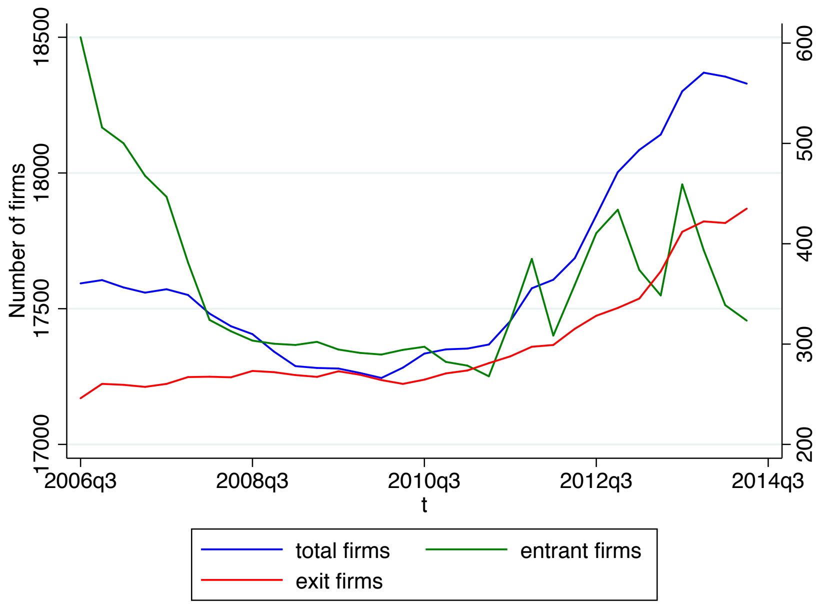


Figure 4: total number of products , total number of entrant products, total number of exit products

This figure plots total number of products , total number of entrant products, total number of exit products each quarter from 2006Q3 to 2014Q2. The left axis is for total number of products, and the right axis is for entrants and exits.

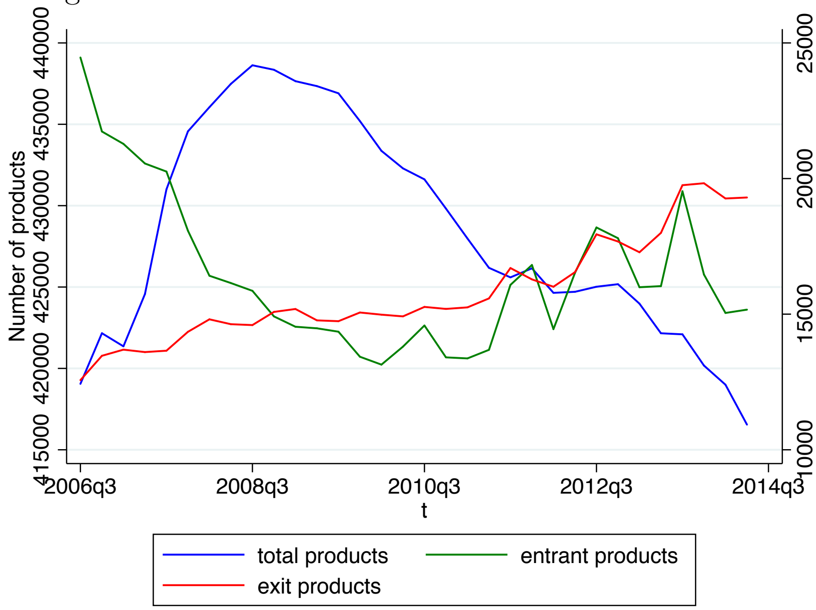
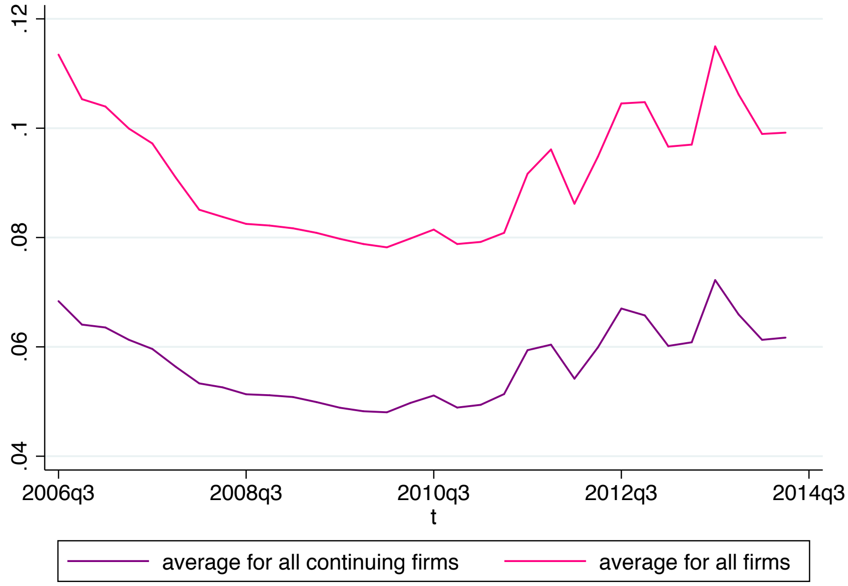


Figure 5: Average reallocation within firm for all firms and continuing firms



A Appendix

A.1 Firm Dataset

Using the product level dataset and the information on the firm, we construct a dataset with the following information for each firm per quarter:

- type of firm Γ_{it} is defined as 1 if it is the first time we see the firm (or all products of the firm are entrants products), -1 if the last period we saw the firms was in t (or all products of the firm are exit products), and 0 otherwise.
- total number of entrant products $N_{it} = \sum_j N_{jit}$
- number of exit products $X_{it} = \sum_j X_{jit}$
- number of continuing products $C_{it} = \sum_j C_{jit}$
- number of all products
  - (i) T_{it} computed as number of observations of products of firm i in t, defined as defined as $N_{it} + C_{it}$
  - (ii) $T2_{it}$ computed as number of observations of products of firm i in $t-1$ plus the changes in the number of products as defined in $\Delta T2_{it}$, i.e. $T2_{it-1} + \Delta T2_{it}$ with $T_{i06Q1} = T2_{i06Q1}$
- change in the number of products
  - (i) ΔT_{it} computed as number of products in t minus number of products in $t-1$

- (ii) $\Delta T2_{it}$ computed as number of entrants in t minus number of exits in t, i.e. change in products is defined as $N_{it} - X_{it}$

- entry rate $n_{it} = N_{it}/T_{it}$ and $n2_{it} = N_{it}/T2_{it}$
- exit rate $x_{it} = X_{it}/T_{it-1}$ and $n2_{it} = N_{it}/T2_{it-1}$
- reallocation within firm $r_{it} = n_{it} + x_{it}$ and $r_{it} = n2_{it} + x2_{it}$
- total revenue of all products rev_{it}
- total revenue of entrant a products $revN_{it}$
- total revenue of entrant b products $revNa_{it}$
- total revenue of entrant c products $revNb_{it}$
- total revenue of entrant d products $revNc_{it}$
- total revenue of entrant d products $revNd_{it}$
- total revenue of entrant e products $revNd_{it}$
- total revenue of exit products $revX_{it}$
- dummy variable to indicate if is a multi-product firm (i.e. if number of products >1) $multi_{it}$
- number of different modules $modules_{it}$
- dummy variable to indicate if is a multi-module firm (i.e. if number of module >1) $multimodules_{it}$
- number of different groups $groups_{it}$
- dummy variable to indicate if is a multi-group firm (i.e. if number of group >1) $multigroups_{it}$
- number of different departments $departments_{it}$
- dummy variable to indicate if is a multi-department firm (i.e. if number of depart. >1) $multidepartments_{it}$
- main module (defined as the module that generates more revenue) $mainmodule_{it}$
- share of revenue in main module $revmainmodule_{it}$
- main group (defined as the group that generates more revenue) $maingroup_{it}$
- share of revenue in main group $revmaingroup_{it}$
- main department (defined as the department that generates more revenue) $maindepartment_{it}$

- share of revenue in main department $revmaindepartment_{it}$
- simple average quality of all products (using the variable *quality0* to *quality7* in the original dataset) $q0_{it}$ $q1_{it}$ $q2_{it}$ $q3_{it}$ $q4_{it}$ $q5_{it}$ $q6_{it}$ $q7_{it}$
- simple average quality of all exit products $q0_{it}$ $q1_{it}$ $q2_{it}$ $q3_{it}$ $q4_{it}$ $q5_{it}$ $q6_{it}$ $q7_{it}$

A.2 Reallocation measures

A.2.1 Aggregate entry, exit and reallocation

We define entry, exit and reallocation as follows

Aggregate entry rate

$$n_t = \frac{\sum_i N_{it}}{\sum_i T_{it}} \tag{11}$$

Aggregate exit rate

$$x_t = \frac{\sum_i X_{it}}{\sum_i T_{it-1}} \tag{12}$$

Aggregate reallocation rate

$$r_t = n_t + x_t \tag{13}$$

Aggregate entry rate by firm net growth of product
(Note that $n_t = n_t^{expanding} + n_t^{contracting} + n_t^{stable}$)

$$n_t^{expanding} = \frac{\sum_i N_{it} 1_{\{i \in \Delta T_{it} > 0\}}}{\sum_i T_{it}} \tag{14}$$

$$n_t^{contracting} = \frac{\sum_i N_{it} 1_{\{i \in \Delta T_{it} < 0\}}}{\sum_i T_{it}} \tag{15}$$

$$n_t^{stable} = \frac{\sum_i N_{it} 1_{\{i \in \Delta T_{it} = 0\}}}{\sum_i T_{it}} \tag{16}$$

We can define the revenue-weighted version as follows

$$\tilde{n}_t^{expanding} = \frac{\sum_i Nrev_{it} 1_{\{i \in \Delta T_{it} > 0\}}}{\sum_i Rev_{it}} \tag{17}$$

$$\tilde{n}_t^{contracting} = \frac{\sum_i Nrev_{it} 1_{\{i \in \Delta T_{it} < 0\}}}{\sum_i Rev_{it}} \tag{18}$$

$$\tilde{n}_t^{stable} = \frac{\sum_i Nrev_{it} 1_{\{i \in \Delta T_{it} = 0\}}}{\sum_i Rev_{it}} \tag{19}$$

Aggregate exit rate by firm net growth of product
(Note that $x_t = x_t^{expanding} + x_t^{contracting} + x_t^{stable}$)

$$x_t^{expanding} = \frac{\sum_i X_{it} 1_{\{i \in \Delta T_{it} > 0\}}}{\sum_i T_{it-1}} \tag{20}$$

$$x_t^{contracting} = \frac{\sum_i X_{it} 1_{\{i \in \Delta T_{it} < 0\}}}{\sum_i T_{it-1}} \tag{21}$$

$$x_t^{stable} = \frac{\sum_i X_{it} 1_{\{i \in \Delta T_{it} = 0\}}}{\sum_i T_{it-1}} \tag{22}$$

We can define the revenue-weighted version as follows

$$\tilde{x}_t^{expanding} = \frac{\sum_i Xrev_{it} 1_{\{i \in \Delta T_{it} > 0\}}}{\sum_i Rev_{it}} \tag{23}$$

$$\tilde{x}_t^{contracting} = \frac{\sum_i Xrev_{it} 1_{\{i \in \Delta T_{it} < 0\}}}{\sum_i Rev_{it}} \tag{24}$$

$$\tilde{x}_t^{stable} = \frac{\sum_i Xrev_{it} 1_{\{i \in \Delta T_{it} = 0\}}}{\sum_i Rev_{it}} \tag{25}$$

Using the definitions above we can define the reallocation counterparts: $r_t^{expanding}$, $r_t^{contracting}$, r_t^{stable}, $\tilde{r}_t^{expanding}$, $\tilde{r}_t^{contracting}$, $\tilde{r}_t^{stable}$

Aggregate entry rate by firm net growth of product and type of firm
(Note that $n_t = n_t^c + n_t^d + x_t^e = n_t^{expanding,entrant} + n_t^{expanding,incumbent} + n_t^{contracting,incumbent} + n_t^{stable,incumbent}$)

$$n_t^{expanding,entrant} = \frac{\sum_i N_{it} 1_{\{i \in \Delta T_{it} > 0 \bigcap i \in \Delta \Gamma_{it} > 0\}}}{\sum_i T_{it}} \tag{26}$$

$$n_t^{expanding,incumbent} = \frac{\sum_i N_{it} 1_{\{i \in \Delta T_{it} > 0 \bigcap i \in \Delta \Gamma_{it} = 0\}}}{\sum_i T_{it}} \tag{27}$$

$$n_t^{contracting,incumbent} = \frac{\sum_i N_{it} 1_{\{i \in \Delta T_{it} < 0 \bigcap i \in \Delta \Gamma_{it} = 0\}}}{\sum_i T_{it}} \tag{28}$$

$$n_t^{stable,incumbent} = \frac{\sum_i N_{it} 1_{\{i \in \Delta T_{it} = 0 \bigcap i \in \Delta \Gamma_{it} = 0\}}}{\sum_i T_{it}} \tag{29}$$

Aggregate exit rate by firm net growth of product and type of firm
(Note that $x_t = x_t^c + x_t^d + x_t^e = x_t^{expanding,incumbent} + x_t^{contracting,incumbent} + x_t^{contracting,exit} + x_t^{stable,incumbent}$)

$$x_t^{expanding,incumbent} = \frac{\sum_i N_{it} 1_{\{i \in \Delta T_{it} > 0 \bigcap i \in \Delta \Gamma_{it} = 0\}}}{\sum_i T_{it-1}} \tag{30}$$

$$x_t^{contracting,incumbent} = \frac{\sum_i N_{it} 1_{\{i \in \Delta T_{it} < 0 \bigcap i \in \Delta \Gamma_{it} = 0\}}}{\sum_i T_{it-1}} \tag{31}$$

$$x_t^{contracting,exit} = \frac{\sum_i N_{it} 1_{\{i \in \Delta T_{it} < 0 \bigcap i \in \Delta \Gamma_{it} < 0\}}}{\sum_i T_{it-1}} \tag{32}$$

$$x_t^{stable,incumbent} = \frac{\sum_i N_{it} 1_{\{i \in \Delta T_{it} = 0 \bigcap i \in \Delta \Gamma_{it} = 0\}}}{\sum_i T_{it-1}} \tag{33}$$

As with the previous definitions of entry and exit, we can define the revenue weighted counterparts. Likewise, using the definitions above we can define the reallocation: $r_t^{expanding,incumbent}$, $r_t^{expanding,entrant}$, $r_t^{contracting,incumbent}$, $r_t^{contracting,exit}$, $r_t^{stable,incumbent}$.

Aggregate entry rate by firm revenue level

$$n_t^{Q(rev)=k} = \frac{\sum_i N_{it} 1_{\{i \in Q(rev)=k\}}}{\sum_i T_{it}}, k = 1, 2, 3, 4 \tag{34}$$

Aggregate exit rate by firm revenue level

$$x_t^{Q(rev)=k} = \frac{\sum_i N_{it} 1_{\{i \in Q(rev)_{it}=k\}}}{\sum_i T_{it}}, k = 1, 2, 3, 4 \tag{35}$$

Using the definitions above we can define the reallocation: $r_t^{Q(rev)=k} = n_t^{Q(rev)=k} + x_t^{Q(rev)=k}$.

Aggregate entry rate by firm multi-product status

$$n_t^{status=k} = \frac{\sum_i N_{it} 1_{\{i \in status_{it}=k\}}}{\sum_i T_{it}}, k = 1, 2, 3, 4, 5 \tag{36}$$

where K=1 if single product firm, K=2 if multi-product but single module firm, K=3 if multi-module but single group firm, K=4 if multi-group but single department firm, and K=5 if multi-department firms.

Aggregate exit rate by firm multi-product status

$$x_t^{status=k} = \frac{\sum_i X_{it} 1_{\{i \in status_{it}=k\}}}{\sum_i T_{it}}, k = 1, 2, 3, 4, 5 \tag{37}$$

where K=1 if single product firm, K=2 if multi-product but single module firm, K=3 if multi-module but single group firm, K=4 if multi-group but single department firm, and K=5 if multi-department firms.

Using the definitions above we can define the reallocation: $r_t^{status=k} = n_t^{status=k} + x_t^{status=k}$.

A.3 Construction of Variables: Compustat

We use the Compustat fundamental annual data from 2006 to 2014. Our sample consists of all firms that have positive data on sales that do not belong to the finance or real estate sectors (SIC classification between 6010 and 6800). To compute total factor productivity (TFP) we follow the methodology of İmrohoroğlu and Tüzel (2014). TFP is estimated using the semi-parametric method by Olley and Pakes (1996). We supplement the Compustat data with output and investment deflators from the Bureau of Economic Analysis and wage data from the Social Security Administration.

We compute R&D as the ratio of research and development expenses (Compustat item XRD) to net sales (Compustat item SALE). We winsorize the variable at the 1% level and use the one-year lagged value of in our regressions. $Size$ is the natural logarithm of total sales. Given that total staff expenses (Compustat item XLR) is mostly missing in our data, we define labor share as the ratio of the number of employees (Compustat item EMP) times the national wage (obtained from the Social Security Administration) over net sales. $StdSale$ is defined as the volatility of annual growth of sales on a quarterly basis winsorized at the 1% level.

We follow Gorodnichenko and Weber (2016) to construct the rest of our control variables. Specifically, Price to cost margin (PCM) is the ratio of net sales minus costs of goods sold (Compustat item COGS) to net sales. Rating (Rat) is the S&P domestic long term issuer credit rating (Compustat item SPLTICRM) where the highest rating category, AAA, is assigned a value of 4.33, decreasing by 1/8 with every rating notch. We use the one-year lagged value of the mean ratings within the year. In order to control for financial constraints, we include the Kaplan-Zingales index (KZ, Kaplan and Zingales (1997)) defined as follows:

$$KZ_{i,t} = 1.002\frac{CF_{i,t}}{AT_{i,t1}}39.368\frac{Div_{i,t}}{AT_{i,t1}}1.315\frac{C_{i,t}}{AT_{i,t1}} + 3.139Lev_{i,t} + 0.283Q_{i,t} \tag{38}$$

where cash flow (CF) is the sum of income before extraordinary items (Compustat item IB) and depreciation and amortization, dividends (Div) are measured as common and preferred dividends (Compustat items DVC and DVP), C is cash and short term investments (Compustat item CHE), leverage (Lev) is the ratio of long term debt and debt in current liabilities (Compustat items DLTT and DLC) to stockholders equity (Compustat item SEQ), and Q is the market value of equity from CRSP as of fiscal year end minus the bookvalue of equity and deferred taxes (Compustat items CEQ and TXDB) to total assets. The first three variables are normalized by lagged total assets. We winsorize all variables at the 1% level before calculating the index and use one-year lagged values of the index in our regressions.