Innovation and growth in the UK pharmaceuticals: the case of product and marketing introductions [†]

Farasat A.S. Bokhari [a,b], Franco Mariuzzo [a,b], Anna Rita Bennato [c]

[a] *School of Economics, University of East Anglia, Norwich NR4 7TJ, UK*
[b] *Centre for Competition Policy, University of East Anglia, Norwich NR4 7TJ, UK*
[c] *School of Business and Economics, Loughborough University, Leicestershire LE11 3TU, UK*

October, 2019

Abstract

New drug introductions are a key to growth for pharmaceutical firms. However not all innovations are the same and they may have differential effects that vary by firm size. We use quarterly sales data on UK pharmaceuticals in a dynamic panel model to estimate the impact of product (new drugs) and marketing (additional pack varieties) innovations within a therapeutic class on a firm's business unit growth. We find that product innovations lead to substantial growth in both the short and long run, whereas a new pack variety only produces short-term effects. The strategies are substitutes but the marginal effects are larger for product innovations relative to additional packs, and the effects are larger for smaller business units. Nonetheless, pack introductions offer a viable short-term growth strategy, especially for small and medium sized businesses.

Key words: Growth; Innovation; Size; Pharmaceuticals; Business unit

JEL Classification: L25, L65, O31, O32

*Corresponding Author: Farasat Bokhari (f.bokhari@uea.ac.uk)

1. Introduction

Theoretical literature suggests a positive relationship between innovation and growth (Geroski, 2005, Aghion and Howitt, 1992) but the empirical literature has reported mixed results with positive links only in some situations, and conditional on firm characteristics (Coad and Rao, 2010, Demirel and Mazzucato, 2012, Deschryvere, 2014, Audretsch et al., 2014). Size, scope, and experience are important factors in determining how much innovative activity a firm undertakes, and whether it results in successful innovation (Acs and Audretsch, 1990, Henderson and Cockburn, 1996). Recent literature has focused on lack of symmetry in the distribution of growth rates and in the returns to innovation to unravel the types of firms that innovate and contribute to growth most, for instance high growth/superstar firms (Coad and Rao, 2008, Capasso et al., 2015, Mazzucato and Parris, 2015). However, not all innovations lead to firm growth, as they may in fact cannibalize sales of existing products (Conner, 1988, Banbury and Mitchell, 1995).

Building on the insights from Hall (1987), Geroski and Machin (1992), Geroski and Toker (1996), Freel (2000), Del Monte and Papagni (2003) and Demirel and Mazzucato (2012), that innovative and/or more R&D active firms outperform non-innovative firms, in this paper we explore the link between innovation, measured by introduction of additional drugs and pack varieties within a therapeutic class, and revenue growth in the UK pharmaceutical sector. The empirical literature that has investigated this relationship has often focused on innovative inputs, such as R&D intensity and patent counts, rather than innovative outputs, such as introduction of new products (see Roper (1997), Flor and Oltra (2004), Coad and Rao (2008, 2010) and Becheikh et al. (2006) for a review). We follow Cucculelli and Ermini (2012) and use a counts based approach of innovative output as a measure of innovation, but highlight the role of two different types of innovations, product and marketing innovations (two of the four types of innovation accounted for by Tavassoli and Karlsson (2015)), as each strategy can have a separate and distinct impact on growth via product differentiation or by price discrimination.[1]

Studies that have used the innovative output approach have often reported lack of a statistically significant positive relationship between growth and innovative output (e.g. Geroski et al., 1997, Bottazzi et al., 2001, Geroski and Mazzucato, 2002, Stam and Wennberg, 2009). However, as pointed out by Corsino and Gabriele (2011), aggregation across heterogeneous products may hide the true relationship, and hence empirical investigations need to focus on sub-markets where products are relatively homogenous, firms draw on a similar knowledge base, and importantly, target the same type of consumers. Following on from that, we use quarterly sales data from the UK pharmaceutical

---

[1]Our distinction between products and packs also relates to heterogeneous measures of innovative output on growth in Akcigit and Kerr (2018) (internal and external innovations) and Caggese (2019) (radical or incremental innovation) and how they differ by firm size.

prescription market during the 2003-2014 period, and measure firms' sales growth within narrowly defined therapeutic classes, which roughly translates into similar patients (customers), and identify the impact of additional drugs and new pack varieties on growth by these subclasses. We call a firm's operation within a narrow therapeutic class a business unit, and by focusing on business units as opposed to firms, we are better able to measure the impact of product innovations (drugs) and marketing innovations (additional pack varieties) on revenue growth by relevant sub-markets.

We follow the empirical literature and estimate reduced form equations where we regress growth on lagged size and counts of products and packs at the business unit level in a dynamic panel model (Evans, 1987a,b). Indeed, revenue and number of products and packs are equilibrium outcomes, and as such, any omitted demand-side factors can induce positive correlation between revenue growth and introduction of new products and packs. Accordingly, in some specifications we treat these measures of products and packs as endogenous, and attempt to find an exogenous source of variation for these variables. We also allow for an interaction between these two variables to understand whether they are complementary or substitutable actions within a business unit.

Our main result is that both the introduction of additional drugs (product innovations) and additional packs (marketing innovations) have a significant impact on revenue growth, and the magnitude is larger for new drugs. New drugs also contribute to the growth in the long run since sales keep increasing, perhaps because more patients switch to the newer formulation. On the other hand, additional packs contribute only to short run growth. Nonetheless, given the price regulation for branded drugs in the UK, a marketing innovation of a new pack may still be a viable short-term business strategy as the cost of introducing a new pack is likely to be much smaller than that of introducing a new drug. This may be particularly helpful for small business units as we find that the impact of additional packs (and products) on growth is much stronger for them. When we further restrict the analysis to a selected sample of surviving business units, we observe that the marginal effect of new products and packs gets smaller, and is closer in magnitude to that of larger business units. This suggests that new introductions by small firms give them a boost in revenue growth, and absent such growth, the lack of innovation also lowers their survival probability. Similar to other studies regarding firm size (e.g. Evans (1987a), Dunne and Hughes (1994) or Bottazzi and Secchi (2005) who find some limited evidence) we observe that smaller business units grow faster than larger ones. Our result that the marginal effect of new products is larger for smaller business units is perhaps consistent with Demirel and Mazzucato (2012), who report positive impact of R&D on growth for small firms, albeit as long as they are patenting consistently.

We make three important contributions to the innovation-growth literature. First, we link innovative outputs to growth (rather than inputs) and explicitly recognize that innovative outputs differ in their appropriability conditions and provide evidence that they have different impact on growth. In

the pharmaceutical sector, drugs with new molecules or formulations are product innovations that are often protected by Europe-wide patents and market exclusivity, and impact revenue growth by channelling demand to better suited drugs or finding new patients. On the other hand, new drug introductions that exploit size or strength variation may be marketing innovations with lower appropriability and affect revenue growth via a different channel, for instance price discrimination. In turn, the impact of these innovations may differ on revenue growth, they are likely to be substitutes, and hence combining these two distinct forms of innovations into one homogenous measure may be inappropriate. The evidence that we provide is likely to be indicative of the pharmaceutical industry more generally outside of the UK as well. Second, we focus on firm sales within narrowly defined therapeutic areas (which we call business units) rather than overall sales or growth by the firm. This is important for studying the link between innovation and growth, as the new introduction of a specific drug may lead to revenue growth within that therapeutic class, but if the firm has operations in many other larger classes, the overall growth at firm level may be small and the correlation may not be picked up by firm level analysis. While our example is from UK pharmaceuticals, where on average a typical firm operates in 9.48 therapeutic classes, the idea carries over to innovative activities of conglomerate firms outside the pharmaceutical industry. Finally, while several papers have already studied growth or innovation in pharmaceuticals, we add to that literature by documenting the innovation-growth link by size of business units, type of innovative output and which one will have a larger impact on growth.[2]

The rest of the paper is organized as follows. The next section describes differences in product and marketing innovations, and sets out the hypotheses we test. Section three describes the data and provides descriptive evidence on the relationship between product and pack introductions and growth. Section four outlines our empirical strategy and discusses sources of endogeneity. Section five gives the results and the last section concludes.

2. Product and marketing innovations and hypotheses

In this paper we use a counts based approach and ask how the introduction of additional products and packs of existing drugs affect the growth of a firm in a therapeutic class (business unit), whether these strategies are substitutes or complements, and whether we should expect results to differ by size of business unit. Firms can introduce new drugs or packs for unilateral reasons if these affect their revenues and/or profits, but may also do so for strategic reasons. Indeed, firms with soon to expire patents can launch additional drugs to deter entry or maintain market shares, or even

---

[2]For some recent papers see, Gambardella et al. (1995), Bottazzi et al. (2001), Orsenigo et al. (2001), Grabowski et al. (2002), Matia et al. (2004), Bottazzi and Secchi (2005), Coad and Rao (2008), Paul et al. (2010), Scannell et al. (2012), Mazzucato and Parris (2015).

increase the prices of their branded drugs post generic entry for their brand loyal market segment (Ellison and Ellison, 2011, Frank and Salkever, 1992, Regan, 2008). Similarly, the originator may make a minor tweak on the original drug, say via a dosage change, as part of a 'product-hopping' strategy (Hemphill and Sampat, 2012, Scott Morton and Kyle, 2011). Doing so can prevent a generic drug from gaining market share.[3] Our focus is on revenue growth, and not alternative reasons for new introductions per se, and hence we discuss below how or why they may influence growth for a business unit.

2.1. **Innovation and growth.** Within the usual four types of innovation (product, process, marketing and organizational), new drugs or packs can be classified as product and marketing innovations. When a firm introduces a new drug within the same therapeutic class it does so by either changing the molecule or the formulation, which typically requires costly and risky R&D activity to receive marketing approval. Instead, introducing an additional strength or pack size requires less medical research (as safety and efficacy of the original molecule/formulation are already established) but it may require further marketing research for placement and pricing strategy. Admittedly, the line between product and marketing innovation is not always clear as a new pack strength can also be seen as product differentiation. Nonetheless, we think of the new drug with an alternative molecule/formulation as product innovation, and new pack as a marketing innovation that allows price discrimination.

The work developed by Geroski and Toker (1996) and Freel (2000), and subsequently by Demirel and Mazzucato (2012), Deschryvere (2014) and Audretsch et al. (2014), among others, show the existence of a positive relationship between innovation and growth at firm level. Based on this literature we test the existence of the same type of relationship between the forms of innovation at business unit (BU) level in the UK pharmaceutical sector.

- Hypothesis 1. Introducing additional products or packs increases sales growth.

2.2. **Returns to product vs marketing innovations.** As discussed in Tavassoli and Karlsson (2015), product and marketing innovations differ in persistence (i.e., repeated events of innovations) due to differences in sunk R&D costs, the 'success breeds success' hypothesis, and appropriability conditions. The authors find that product innovations are more persistent than marketing innovations due to differences in (inter alia) appropriability conditions. In turn, this implies that increase in revenue (or growth) associated with new products will be larger than with new packs. Certainly,

---

[3]This is because the generic would have filed for, and gained entry for a different dosage, and patients may have been switched to the newer formulation via marketing or other mechanisms to curtail demand for the original formulation. Alternatively, follow-on drugs (by originators or new entrants) may provide substantial welfare gains, perhaps by lowering the side effects of the pioneer drugs, changing the delivery mechanism, or targeting a new sub-population and effectively increasing the market size (Arcidiacono et al., 2013, Bokhari and Fournier, 2013).

the two types of innovations we discuss here differ in appropriability conditions since a new drug introduction can often be protected via patent or other marketing exclusivity rules, but the same IP protection is not necessarily extended to pack innovations. In introducing a new drug, they may also meet some pent-up demand to serve new patients, or alternatively allow low valuation patients to switch to the new formulation if they value the alternative more for medical reasons.

For example, in 2001, Shire introduced Adderall XR as a once-a-day extended release formulation of its original drug called Adderall, which was to be taken multiple times a day to treat ADHD symptoms (for school aged children, having an option to give the medication once a day is an important innovation due to difficulties in administering a drug during school hours). In 2000, Shire had 31.1% of the US ADHD drug market. At the time, there were several other firms also producing ADHD drugs using other molecules, but due to a patent that was valid until 2002, Shire was the only one using the specific molecule used in Adderall. When Shire introduced the XR version, it acquired a new patent on the formulation and obtained a market exclusivity for the XR version of that specific molecule. By 2003, share of the generic Adderall by competitors was 7.6%, that of the original Adderall by Shire was 2.9%, and of new Adderall XR was 23.8% (Bokhari and Fournier, 2013). As a result of this strategy, Shire's sales for these products increased from $296.8 million in 2000 to $624.35 million in 2003, a 110.36% increase in revenue.

TABLE 1. Price per mg as function of packs

	Add Pack	Remove Pack	Time (T)	Time Sq (T^2)	Time Cb (T^3)	Constant
(ln) price/mg	0.053^a	-0.002	0.007	$-7.3e\text{-}05^b$	$1.85e\text{-}07^b$	1.702^a
	(0.011)	(0.008)	(0.005)	(3.34e-05)	(7.96e-08)	(0.229)

Fixed effects regression (with individual drug ids) for log price of drug per mg on adding and removing packs within a drug. Total 152,673 observations from monthly prices over 10 years for drugs in tablet or capsule form. Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively and standard errors are clustered by therapeutic area of the drug.

Alternatively, firms can change price per unit for the same drug by introducing a new pack, which can vary either by strength or size (or both). For instance, in May 2000, 3M's Aerolin (a salbutamol used for treating bronchial asthma) was available in the UK for £12 per pack of 56 pills of strength 8mg. Two months later, a 56 pills pack of 4mg strength was also available for £10.55, so while the price per pill and per pack decreased, the price per mg nearly doubled from 2.67 to 4.71 pence. An example of packs varying by size is Abbot's Prothiaden, an SSRI with dosulepin hydrochloride as the main active ingredient for treating depression was sold in the UK in 2003 as a 75mg 28-tab pack for £4.20 per pack or as 75mg 500-tab pack for £85.81, which works out to £2.00 per gram and £2.88 per gram respectively. Indeed, there are many examples of non-linear pricing in the UK pharmaceutical data. Our data has several similar examples which indicate that among branded

drugs, the introduction of a new pack is followed by an increase of approximately 5.3% in price per gram for tablets or capsules (see Table 1).[4]

Thus, due to the difference in appropriability conditions (new patients and marketing exclusivity) between products and packs, we state our second hypothesis.

- Hypothesis 2. Revenue growth is higher for new products relative to new packs.

2.3. **Substitutability between product and marketing innovations.** If launching additional drugs helps the business unit create product differentiation (so patients select into the drug type most medically suitable for them for which they or NHS pays a higher price) while an additional pack size helps in creating price discrimination, these two strategies may not be complementary but rather work as substitutes. In Appendix A.1 we provide an example which shows launching an additional drug or a new pack increases firm revenue via product differentiation or by price discrimination, but doing both simultaneously may not be optimal. The intuition is that with product differentiation, the market is segmented and low value consumers switch to the new product if they value it more. If there is also price discrimination in the original product (due to an additional pack being available for the original drug), some low value consumers may not switch to the new product since they are already getting the original product at their low valuation (market is covered), in which case implementing both strategies simultaneously may not be optimal relatively to implementing them separately.[5] Thus, we have the following hypothesis.

- Hypothesis 3. Introduction of new products and packs are substitute strategies for revenue growth.

2.4. **Returns by size.** While we have made a distinction between product and marketing innovations, they can further be classified as being radically different from existing product lines or not. For instance, launching an additional extended release tablet when a regular tablet is available versus launching a transdermal patch may be a more innovative drug introduction. Caggese (2019)

---

[4]While we consider a new pack to be a marketing innovation that allows some form of price discrimination, indeed it can also be considered a form of product differentiation, for instance where a patient cannot cut a capsule in two parts to take half-strength drug even though that may be medically ideal for them.

[5]In the context of NHS where patients do not directly pay for prescription drugs except for a small co-pay, it is still possible to talk about a 'low value' consumer. In the UK after a drug is approved for marketing authorization, the National Institute of Clinical Excellence (NICE) undertakes cost-effectiveness (CE) of the new treatment option. The CE analysis is based with a target patient population in mind, where the patient population is defined based on demographics, overall health condition, and reported symptoms. The CE analysis then attempts to measure changes in Quality Adjusted Life Years (QALYs) associated with the new therapy relative to the status quo, and may find one group of patients for whom the new drug is not cost effective while for another it is. This would then translate into 'recommendation with restriction' and also permeate in clinical guidelines describing for which group of patients the new drug is recommended and for whom it is not. See the next section 'Background and data' for more information on role of NICE.

calls these radical and incremental innovations (a relabelling of Schumpeter's concept of 'disruptive innovation') and states that if there are no financial frictions the firms that rapidly grow in productivity and size are the young and small firms as they tend to prefer radical innovations more than large firms. On the other hand, financial frictions reduce competition in the market and the by-product is that young and small firms find less profitable the investment in radical innovations, making incremental innovation more appealing. Similarly, Akcigit (2010) shows that innovation activities (measured by R&D intensity) and high-quality innovation (measured by patent citations) relate inversely to firm size, and Akcigit and Kerr (2018) show that small firms contribute most to the relative rate of radical innovation. If indeed smaller BUs introduce more radical/innovative drugs compared to larger ones, then we would expect that returns from innovation will be higher for smaller BUs, which we test as a hypothesis.

- Hypothesis 4. Small business units have a higher revenue growth from innovation than larger business units.

We test these hypotheses using revenue growth equations for a business unit as a function of new products, packs and their interactions.

3. Background and data

To bring either a new drug or a generic copy to the market in Europe, a pharmaceutical firm requires market authorization (MA) from a national authority (Medicines and Healthcare products Regulatory Agency (MHRA) in the UK) or, as of 1995, from the European Medicines Agency (EMA). The process starts with the firm filing for a new drug application (in case of a new molecular entity) or an 'abridged' application for generic entry. In the former case, MA is granted after establishing safety and efficacy via three phase clinical trials that take several years to complete while in the latter case, the generic applicant references data for an earlier drug and aims to establish bioequivalence to it.

Drug development is expensive and risky as not all new discoveries make it to the market in the form of a new drug (Torjesen, 2015, DiMasi et al., 2016). Further, because patent life is 20 years from initial file date, and significant time is lost in drug development, EU provides two routes that allow innovators to extend the exclusive marketing of their products. The first, available since 1992, is the Supplementary Protection Certificate (SPC) which allows originators to extend the patent for up to five years after the expiration of the original patent, or fifteen years from the first marketing authorization in the EU, whichever is less. Second, there is an explicit data exclusivity period which was introduced in 1984 at the EU level (prior to that, drug approval was at the national level and with varying rules) during which a generic firm may not reference the originators data. Initially,

data exclusivity extended either to six years or ten years from the start of MA date depending on the member state (UK had 10 years). In 2005, a new '8+2(+1)' exclusivity period was introduced at the EU level which provided unified rules of exclusivity across all member states – eight years of data exclusivity during which a generic cannot file for an abridged application, plus two additional years of market exclusivity, i.e., the generic may file the abridged application, but not market the drug, and a final one additional year of market exclusivity for new indication(s) if they constitute a significant clinical benefit.

In the UK there is an additional step which is particularly relevant for new expensive drugs. National Institute of Clinical Excellence (NICE), which was established in 1999, is charged with undertaking health technology appraisals including those for new medicines as well as providing clinical guidelines for physicians. While NICE is not part of EMA and does not assess safety and efficacy of drugs, it undertakes cost-effectiveness of EMA/MHRA licensed drugs relative to existing practice in the NHS.[6] Not all new drugs are necessarily reviewed by NICE (Ward et al. (2014) identified 134 new MAs by EMA between 2005 and 2011 and found that only about 54% were apprised by NICE). However if apprised, a positive review means that the drug must be covered by NHS while a negative review means that it will not be covered by NHS (in other cases NICE may give "recommend with restriction" for a smaller population relative to the one listed in original drug approval application). As reported in Jaska et al. (2014a,b), between 2007-2013, the overall negative recommendation rate was 32% but this was mostly driven by oncology related drugs which are often used in hospitals (52% for oncology vs 16% for non-oncology reviews). Nonetheless, drugs that are not eligible to be covered by the NHS are either available over-the-counter (OTC) or face a thin online market. The OTC drugs have reached £2.62bn in Britain in 2016, versus £17.4 billion NHS expenditure in England in 2016/17 (Connelly, 2017, NHS Digital, 2018).

Use of UK data is attractive, as price is less regulated compared to some of the other EU member states, and because in many member states, reimbursement is often in terms of price per defined daily dose, which lowers the potential for (second-degree) price discrimination via pack variations. Additionally, prices from the UK are often used as comparison when setting prices in other European countries.[7] Unlike some other European countries that use some form of direct price control, the UK instead employs indirect methods and regulates profit on sales of branded drugs to the National Health Service (NHS) under its Pharmaceutical Price Regulation Scheme (PPRS). The scheme,

---

[6]NICE provides appraisals for England and Wales, while the Scottish Medicines Consortium (SMC) plays a similar role in Scotland including advice on drugs apprised through NICE. Further, drugs that cost NHS over £20m a year in the first three years of launch require a commercial discussion between NHS and the manufacturer.

[7]A common tool is the external reference pricing, employed for instance in France, Greece, Ireland, Italy, the Netherlands and Portugal, where the relevant health authority uses the average price from other EU countries to set the maximum reimbursement rate for a branded drug and UK prices often feature in that reference pricing (Kanavos, 2003, Ruggeri and Nolte, 2013).

which was initially introduced in 1957, has evolved over the years, but as of 1986 it applies only to branded drugs. The terms of this scheme are revised approximately every five years. Under the PPRS, firms are free to set prices of branded drugs, which the NHS will reimburse, but prices are indirectly controlled through caps on the overall return on capital for research active firms. Manufacturers can set the price of new drugs without pre-approval by the Department of Health (DH). However, price increases for existing branded drugs need to be justified and approved by the DH, and in practice may even require a reduction of price of some other drugs in the firm's portfolio to justify the increase in price elsewhere.

We use the 2003-2013 British Pharmaceutical Index (BPI) data series by Intercontinental Marketing Services (IMS), a data set which provides national level monthly sales at the package level for all drugs sold in the UK. The BPI data set contains information in terms of total shipments by nominal sales value and various measures of quantity from wholesalers to retail pharmacies and dispensing doctors, but does not include direct sales from manufacturers to hospitals or to non-pharmacy stores (e.g. grocery stores). Individual drugs in the data are identified by manufacturer, product name, and pack variation. The data also includes information on main/active molecule(s), strength and form, as well as if the drug is branded or generic, over the counter or prescription, and if it is reimbursable by NHS or not. We restrict our analysis to prescription drugs covered by the NHS, i.e. over-the-counter (OTC) and non-reimbursable drugs are excluded as these would be sold outside of pharmacies as well, and we do not have complete sales data on those drugs. In 2005, the OTC or self-medication market represents about 12.03% of all pharmaceutical sales in the UK (Habl et al., 2006, p.707). In a later section we also comment on how our results change if we do not exclude these drugs from the analysis.

For each individual item at the pack level, the data set lists the associated four-digit anatomical therapeutic chemical code (ATC4) and three-digit new form classification code (NFC3). The ATC codes allow drugs to be classified by active ingredients and are further refined by anatomical, therapeutic, pharmacological and chemical subgroups, while the NFC codes provide information on various forms of the drug, e.g., tablet, capsule, extended release, liquid, cream etc. In our analysis, we use the combination of four-digit ATC classification and firm identity to define a 'business unit', and NFC codes to differentiate among various products within a business unit. We aggregate monthly sales data from individual drugs at the pack level to ATC4 classification by manufacturer and quarter. For example, acid pump inhibitors (A02B2) and psychostimulants (N06B0) produced by Novartis are treated as two separate business units (BUs). The final sample consists of sales data from 218 pharmaceutical firms operating in 385 different ATC4 classes spanning 2,090 business units (i.e., firm-class pairs) observed over 40 quarters (Q2 2003 to Q1 2013) for a total of 56,070

observations.[8] All sales figures are adjusted for inflation using the consumer price index as deflator with base period set to second quarter of 2003.

UK log sales by business unit, period t vs t-1. All numbers only for Rx market.
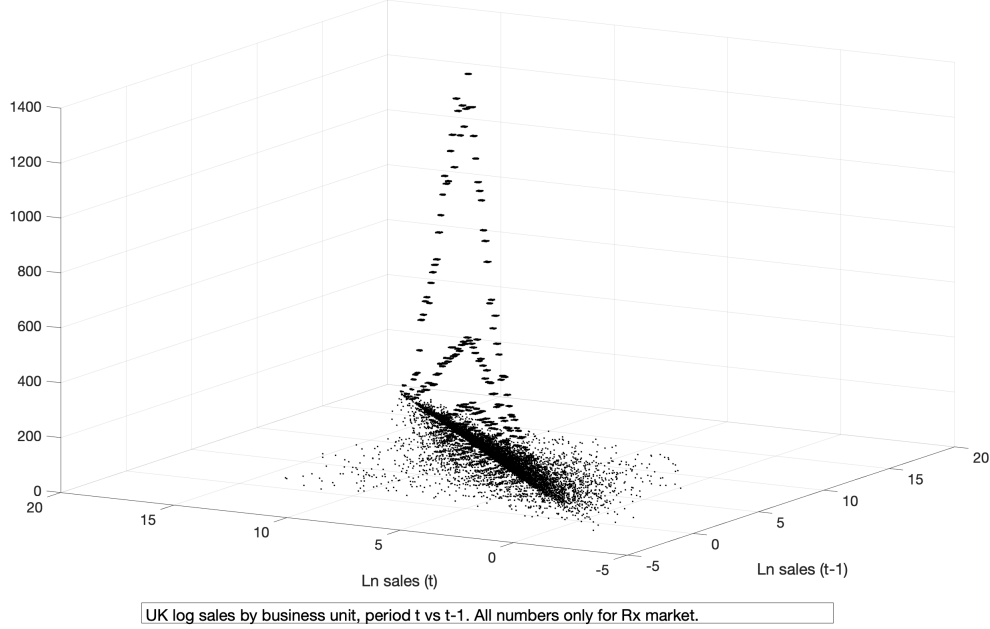
Figure 1. Log-log scatter plot of revenues

As might be expected, sales are highly skewed and we use log of sales (distribution of log sales is relatively symmetric and is given in the appendix in Figure A-1). Based on tertile distribution of the initial log size of a BU (i.e., first period), we classify them as small, medium and large and where the cut-off points are 9.84 and 12.50 in log size.[9] In terms of growth, defined as the change in log sales, we find some variation by the size of the business unit. Figure 1 plots the natural logarithm of sales in period t against the one period lagged value of the same variable (each period is a quarter), and plots the frequency of such pairs on the z-axis. Values along the 45-degree line indicate absence of growth for the BU while those above the line show positive growth. Variation appears to be larger for middle level of sales, and on average, small sized BUs appear to grow more than very large sized ones. Nonetheless, there are relatively fewer observations that deviate from the 45-degree line and most are very near the line.

---

[8]The BPI does not separately identify all generic manufacturers and hence within each ATC4 class multiple generic manufacturers are treated as one firm. As a robustness check, we dropped all observations relating to these unidentified generic manufacturers and discuss them in Section (5.8).

[9]A small BU corresponds to an average firm size of £80.5M, a medium BU corresponds to average firm size of £80.7M, and a large BU corresponds to an average firm size of £138M.

We define a product as a unique combination of proprietary or international nonproprietary name listed in the database, its specific formulation (tablet, extended release, liquid, gel, etc.), the associated ATC4 therapeutic class, and the name of the manufacturer. If any of these change, we count it as different product. Next, for a given product, packs can vary either by size (28 pill pack vs 14 pill pack) or by strength (100MG vs 250MG), and after counting total number of packs per product, we define pack variety (or just variety for short) as the total number of packs for a BU minus the total number of products. Thus, at the BU level,

$$\text{pack variety} = \#\text{packs} - \#\text{products} \tag{1}$$

and our pack variety measure begins at zero when each product is available in only one pack size. This is so that when a new product is introduced, it does not lead to an automatic increase in pack varieties offered by the BU as well (since each product must come in at least one pack size).

While we use count of drugs in business unit to measure innovation, a shortcoming of this approach is that it weighs all new drugs equally, even though some drugs may be small tweaks on the original variant as mentioned earlier. In our data, it is not possible to construct measures that capture the innovativeness of new drug. For instance, studies that use patent counts try to overcome this difficulty by using citation count of a patent.

Merger and acquisition activities during the period are handled by IMS by retroactively reassigning the sales and associated products to the end-of-period corporation that owns them as if they owned it for the entire period. Thus, if two firms merged during the observed period they appear to be a single firm from the start (a similar method is used for instance in Bottazzi and Secchi (2005)) and hence a merger would not lead to an increase in the count of drugs and revenues due to the merger. Generally, a similar rule applies to product line acquisitions, however here we found some inconsistencies in data which lead to small errors in counts of products. Thus while we use count of drugs as a measure of innovative activity, this is not a perfect measure but based on the sample of cases that we investigated this is not a frequent issue.[10]

Figure 2 plots the total sales and average number of product and pack varieties per BU for all ATC4 classes. Consistent with other industry reports of a slow down in manufacturing, our data shows that revenues dropped by 22 percent during the 2004-2012 period. The figure shows that total sales declined over the period, but also that there is some seasonality in sales. More interestingly,

---

[10]For instance, in a few cases of single or a small product line acquisitions, we found that both the product name and company name changed in the raw data. Thus, the product in question appears with the initial firm up to the acquisition date, and then with the acquiring firm along with a modified product name after that date. This implies that in these cases, the count of products decreases by one for the selling firm and increases by the same amount for the acquiring firm without there being any innovation. Since a separate exhaustive list of all such acquisitions is not available, we cannot make these corrections systematically.

UK Sales by Business Unit. Sales £ in Billions and in constant value (2003-Q1).
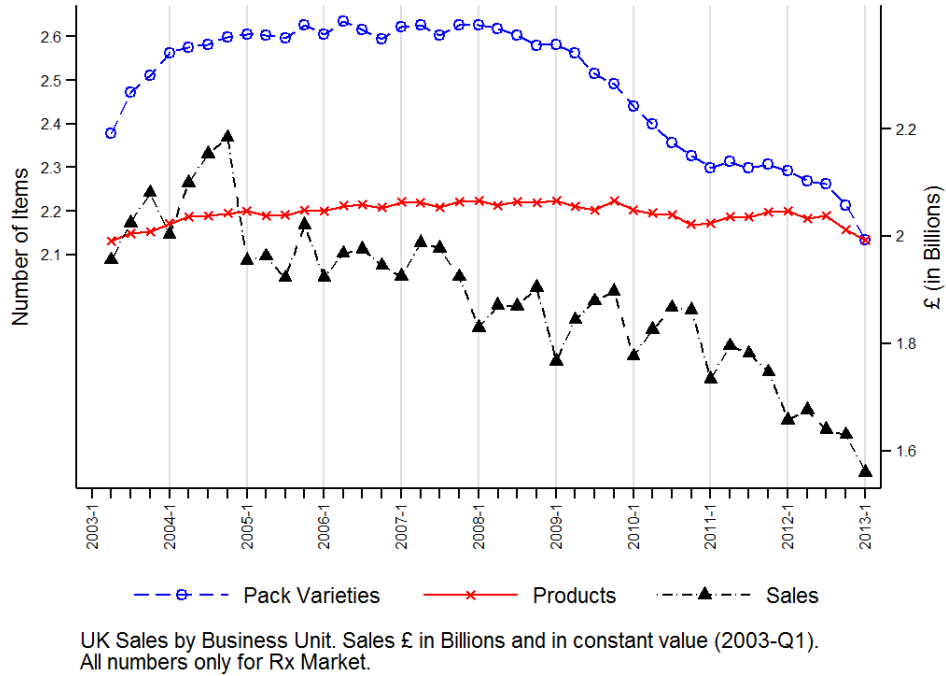All numbers only for Rx Market.

FIGURE 2. Products, pack varieties per BU and total sales

while the average number of products per BU is stable around 2.2, there is a decline in the average number of pack varieties starting with 2009 and continues for the next two years until it settles to a new steady state value. While this decline in the number of pack varieties coincides with the global financial crisis, we do not observe a similar decline in pack varieties in the OTC and non-NHS reimbursable drugs (recall we also have data on those drugs that is not included in this analysis). The start of the decline however does coincide with the last PPRS update in the observation period, which was negotiated at the end of 2008 and was effective for the 2009-2013 period. The figure suggests that there is a positive relationship between aggregate sales and average products/varieties per BU, and that perhaps sales are in decline in the UK since firms do not bother introducing new products and packs. This however is not necessarily true. Similar figures drawn at class level indicate different relationships between sales and product and pack introductions at this disaggregated level (see Figure A-2 in the appendix for some selected classes).[11]

Finally, and before analyzing growth more systematically in the next section, Table 2 shows growth just before and after a BU introduces a product or a pack, and compares it to the growth of

---

[11]Figure A-2 in the appendix shows that sales, products and packs are increasing in the second panel (top right, for stomatologicals) where increase in pack varieties is followed by an increase in sales. By contrast, the middle two panels for seroton and vitamin D show an increase followed by a decrease in sales, and do not exhibit a positive correlation with products and packs over the entire period. Similarly, the bottom two panels (gastroprokinetics and stimulant laxatives) show very little change in products and pack varieties and yet there is considerable change in sales, showing that at the micro level, it is not true that a decline in sales is due to firms not introducing products and packs.

competitors in the same ATC class (the table reports four period moving average of growth before and after introductions). On a class-by-class basis, we find that when a BU introduces a new pack, on average its growth changes from 0.0319 to 0.0507, and for the competitors in the same class but not introducing in that period, growth changes from -0.0103 to -0.0243 (these figures are statistically different from zero). The impact of product introductions has a similar pattern and is given in Table 2.

TABLE 2. Moving average (order four) of revenue growth

	New Packs		New Products	
	Before	After	Before	After
[†]Introducing BUs	0.0319^a	0.0507^a	0.0079	0.0785^a
(Std.Err.)	(0.0073)	(0.0071)	(0.0091)	(0.0110)
[‡]Non-Introducing BUs	-0.0103^a	-0.0243^a	-0.0120^a	-0.0225^a
(Std.Err.)	(0.0026)	(0.0028)	(0.0030)	(0.0033)

[†]Four-quarter average growth for BUs introducing a pack or product in a given period.
[‡]Four-quarter average growth of non-introducing BUs in the class and period where a pack or product was introduced.
[a]Significant at 1% level.

4. EMPIRICAL SPECIFICATION

As our point of entry, we adopt the empirical firm growth model first introduced by Evans (1987a,b) which specifies the change in log size as a linear function of lagged log size, age, other firm specific characteristics of interest, and an additive error term. However, we do so in the context of business units and panel data. Specifically, let $g_{bt} \equiv (\ln R_{bt} - \ln R_{bt-1})$ denote the revenue growth for business unit b (composed of firm f in anatomical class c) between period t and $t-1$. We model the growth equation by an autoregressive distributed lag specification with M lags of the dependent variable and up to L lags of the time varying strategic variables (products and varieties) as

$$g_{bt} = \sum_{m=1}^{M} \gamma_m g_{b(t-m)} + \sum_{l=0}^{L} \left(\beta_{1l} p_{b(t-l)} + \beta_{2l} v_{b(t-l)} + \beta_{3l} p_{b(t-l)} v_{b(t-l)} \right)$$
$$+ \theta_1 s_{b(t-1)} + \mathbf{x}_{bt} \boldsymbol{\theta}_2 + \mathbf{x}_{ct} \boldsymbol{\theta}_3 + u_t + \alpha_b + \varepsilon_{bt}, \quad t = L+1, \ldots, T, \quad L \geq M.$$
(2)

In the equation above the business unit growth in period t is a linear function of lagged growth (with M lags), current and past values of number of products and pack varieties per product (p, v and their interactions up to L lags), and where the latter variables are potentially endogenous. In the final specification we set $M = 1$ and $L = 4$. The specification also includes other relevant variables, prominently the one period lagged size ($s_{b,t-1}$), which we capture with the lagged natural log of revenue. Note than in addition to measuring the impact of size on growth, earlier literature

has also highlighted the relationship between size and innovative output (Acs and Audretsch, 1988, Graves and Langowitz, 1993). Thus, not including size would cause an omitted variable bias as it would be correlated with the variables of interest, i.e. counts of products and packs.

We also include other class-time varying variables such as the number of firms, Herfindahl index, and lagged values of number of products and pack varieties at the class level in the vector $\mathbf{x}_{ct}$, and incorporate additional exogenous controls that vary at the BU level in $\mathbf{x}_{bt}$ (specifically dummy variables to indicate if the business unit has sales due to generic products and if its products are available via parallel imports in the UK and lagged value of mean drug price in log form).[12] All variables listed here are described in full detail in the appendix in Table A-1 and correlations between them are given in Table A-2. Finally, note that while the model includes an interaction term and up to four lags of this term ($p_{b(t-l)}v_{b(t-l)}$ where $l = 1, 2, 3, 4$), we do not include any mixed lags (for instance, two lags of products interacted with three lags of pack varieties). In part, this decision is driven by the desire to keep the model parsimonious, but it is also based on low correlations between product and pack introductions in different periods. For instance, while the correlation in change in number of products and packs from same the lags is fairly high (between .26 and .27, and hence these terms are included in the model), the correlation between mixed lags is typically .05 or less (a full matrix showing correlations in change in products and packs is given in Table A-3 in the appendix).

For short panels it is common to let the time effect be fixed, here given by u_t and hence we exclude it from the composite error term, which is given by the sum of the business unit unobserved heterogeneity, α_b, and the pure idiosyncratic error term, ε_{bt}. We assume the idiosyncratic error to be serially uncorrelated and uncorrelated across ATC4 classes. Because of the presence of a lagged dependent variable we follow the dynamic linear panel model literature and take a first difference of equation (2) to remove inconsistencies of the parameters associated to the various lags of the dependent variable. Thus we estimate the growth equation above in first difference form given by

$$
\Delta g_{bt} = \sum_{m=1}^{M} \gamma_m \Delta g_{b(t-m)} + \sum_{l=0}^{L} \left[\beta_{1l} \Delta p_{b(t-l)} + \beta_{2l} \Delta v_{b(t-l)} + \beta_{3l} \Delta \left(p_{b(t-l)} v_{b(t-l)} \right) \right]
$$
$$
+ \theta_1 \Delta s_{b(t-1)} + \Delta \mathbf{x}_{bt} \boldsymbol{\theta}_2 + \Delta \mathbf{x}_{ct} \boldsymbol{\theta}_3 + \tau_t + \Delta \varepsilon_{bt}.
\tag{3}
$$

While the first difference form eliminates the unobserved heterogeneity specified earlier as α_b term, it is worth noting that by using a first difference estimator we also rule out estimating the effect of any time-invariant factors such as the age of the business unit, which may be important (see for instance Cabral and Mata, 2003, Coad et al., 2016).

---

[12]Our raw data comes with a price of a 'counting unit', which roughly translates into the smallest unit. For instance, for solid form drugs, a counting unit is a tablet. We use this measure of price to compute share weighted average price per counting unit over all drugs sold by the BU, where shares are based on revenues.

Endogeneity. New products and packs are not randomly launched and there may be potentially multiple sources of endogeneity. The first source is any omitted variables that are correlated with growth as well as with product and pack varieties (demand side factors may bring more products/packs and sales growth). de Frutos et al. (2013) document that better drugs are advertised more intensively, and advertising has a positive effect on prices. In turn, this could imply that the magnitude of the positive relationship between launches and revenues may in part be driven by the size of the marketing/sales force, which can easily mention the launch of a new product/pack.[13] However, if the sales force is roughly proportional to the size of the BU and the proportionality does not change from one period to the next, our first differencing approach should account for this missing variable. If however the proportionality does change, but is still proportional to the size of the BU, inclusion of the lagged size should act as a reasonable proxy (we also included the lagged size of the firm rather than just that of the BU and it did not change our results in any meaningful way). Thus, our first remedy is to use a first difference estimation, which should account for factors not changing over time but affecting both revenue sales and new launches.

However, the first difference equation cannot be estimated consistently by OLS since now $\Delta \varepsilon_{bt} = \varepsilon_{bt} - \varepsilon_{b(t-1)}$ is correlated with $\Delta g_{b(t-1)} = g_{b(t-1)} - g_{b(t-2)}$ because $g_{b(t-1)}$ depends on $\varepsilon_{b(t-1)}$ per the initial specification. To this end we use the time series dimension of the panel and obtain instruments for the change in the lagged dependent variable by following Anderson and Hsiao (1982). Specifically, for $m = 1$, we can instrument $\Delta g_{b(t-1)}$ with $g_{b(t-2)}$ since the latter term is uncorrelated with $\Delta \varepsilon_{bt}$ by construction, and for $m > 1$, Anderson and Hsiao propose other lagged dependent variables as instruments for themselves, i.e., that they can be treated as exogenous.

A second potential source of endogeneity is due to sample selection associated with survival. As derived by Jovanovic's model of firm efficiency, and verified in many empirical models, survival will increase with age and size of the firm (Jovanovic, 1982, Dunne et al., 1988, Dunne and Hughes, 1994, Phillips and Kirchhoff, 1989, Baldwin and Gorecki, 1991, Audretsch and Mahmood, 1994). Alternatively, it may be that new small firms grow slowly with age, not necessarily due to cost differences, but because of differences in demand side fundamentals such as consumer experience with newer products, and they catch-up if they initially survive the lack of demand for their products. Either way, it may be true that small and innovative start-up firms grow faster, and if they

---

[13]Pharmaceutical firms use various marketing instruments such as giving free samples, "detailing" physicians, advertising in professional magazines, sponsoring at conferences, and in the case of US and New Zealand, also engage in direct-to-consumer advertising. While it is well established that the overall marketing expenditures are high in this industry (see Hurwitz and Caves, 1988, Manchanda and Honka, 2005), there is also some documentation over how they change over life cycle of a product and with new introductions. Huskamp et al. (2008) find that firms facing imminent generic entry that launch a new formulation shift promotional spends from the original drug to the new formulation, and do so prior to generic entry. They also report that the overall promotional expenditures for the molecule decrease after generic entry takes place. Landsman et al. (2003, p.206) also report that firms spend largest portion of marketing budget on physician detailing at the time of a new launch.

do not launch new products/packs, they do not survive. Then if we were observing only surviving BUs, our coefficients would be biased downwards. Fortunately our data does not suffer from this selection problem as we observe all BUs, and not just those that survive the entire duration. In fact when we restrict the analysis to only surviving BUs, the sample looks more like the established bigger BUs, and the estimated marginal effects move in the anticipated direction.

Finally, the third source of endogenity is the potential reverse causality, where BU's that grow more are the ones that introduce more products and packs (Coad and Rao, 2010). In turn, this implies that $E(\Delta p_{b(t-l)}, \Delta \varepsilon_{bt}) \neq 0$ and $E(\Delta v_{b(t-l)}, \Delta \varepsilon_{bt}) \neq 0$, and obviously $E(\Delta (p_{b(t-l)} \times v_{b(t-l)}), \Delta \varepsilon_{bt}) \neq 0$, with $l = \{0, 1\}$ in the original levels equation. Thus, we need to treat Δp_{bt}, Δv_{bt}, their interaction term, and in some specifications their lagged values as endogenous variables (in addition to $\Delta g_{b(t-1)}$ already discussed above).

To address these concerns, we use instrumental variables which are constructed as the average number of products and packs in other closely related business units of the same firm (defined as counts in those ATC4s that are within the same ATC2 level classification), along with their interactions and lagged values. Details on constructing the exact variables and variation in them are in the Appendix A.3. For these additional instruments to work however, two conditions must hold. First, a firm's propensity/ability to introduce additional products and varieties is driven by factors that are common across its various BUs, and hence the number of products and varieties are correlated across BUs. These correlations could arise due to firm level access to the same capital markets, common large chemical libraries from which they draw their search for a new product (see Thomke and Kuemmerle, 2002), common legal and product development departments dealing with market approval and other regulatory needs.

The second assumption (giving exclusion restriction) is that any changes in the number of products and varieties in one BU do not directly affect growth in another BU. For instance, boost in sales in another unit due to introduction of new products/varieties is not used to cross-subsidize marketing efforts in the reference BU which in turn would affect growth. Our instruments would also fail the exclusion restriction if there is significant brand or firm loyalty among prescribers *across* therapeutic classes. For example, if a new drug is introduced by a BU and is indeed a great success, it generates brand loyalty so that patients also start purchasing drugs by the same firm in other therapeutic classes as well. Indeed Bernard et al. (2018) show that when there are demand-scope complementarities, an increase in a firm's total product scope can increase demand for all of the firm's other products, and find evidence from Belgium manufacturers exporting own and sourced products. However, given that we focus on prescription drugs and not over-the-counter medication, we assume that physicians are guided by prescribing rules within NHS, and are strongly encouraged to prescribe by international non-proprietary name of the drug rather than by proprietary

brand/firm name (see for instance Rawlins, 2015), and hence may be less susceptible to brand loyalty across therapeutic classes.[14] In fact, in many cases it may not even be the same physician who prescribes medication for different health issues across therapeutic classes: a specialist dealing with children's ADHD and prescribing drugs for that treatment may not have any similar experience or exposure to successful drugs launched by the same firm in a different class. The more distinct the therapeutic classes, the less likely the risk of brand loyalty, but then the weaker may be the instrument as well (we discuss this further in the appendix).

However, a more serious limitation of our instruments is that our strategy does not work for firms with single BUs. Restricting the analysis to firms with more than one BU may introduce other selection biases, while constructing instruments from more aggregated ATC2 classes can pick up effects driven by other factors. Thus, we provide two sets of results: by restricting the analysis to only those BUs that operate in more than one ATC4 class, but also on the full sample and ignoring the third potential source of endogeneity. As we discuss later, the results from the two cases are not very different and hence we prefer the analysis with the full sample. Finally, we provide additional evidence for the first assumption above via first-stage F-tests, and while the second assumption cannot be directly verified, we show over-identification tests as suggestive evidence in the Online Appendix (in Table B-2 and Table B-3), and recognize that there is no true test for IV assumptions.

5. Results and discussion

5.1. **Descriptive statistics.** Table 3 provides summary statistics for all the main variables of interest in our empirical model. Consistent with sales over time shown in Figure 2, the average BU growth over the ten-year period is -0.02, while the average size of a business unit is 11.26 on the log sales scale (corresponding to £1.34 million per quarter). However, there is significant variation in both growth and size of BUs cross-sectionally (*between* BUs) as well as over time (*within* BUs) but for growth, the variation is larger over time than cross-sectionally (within SD is 0.71 and between SD is 0.42). In fact, outside of growth itself, between variation is larger for all variables of interest relative to variation over time. In terms of the structure, while the average number of products and varieties per BU are 2.19 and 2.49, the same two variables at the class level are 9.03 and 10.15, implying that there must be roughly three additional firms with an average of just above two products, each competing with any given BU. The lower part of the table confirms this and shows that on average there are 4.23 firms per ATC4 class. The inter-class competition ranges from a monopoly to up to 29 different firms offering their products in the same class of drugs. In terms of the Herfindahl index, which ranges from 0.13 to 1.00 with an average value of 0.69

---

[14]"... *the high rates of generic prescribing in the UK owe much to the fact that general practitioner prescribing systems automatically convert brand names to generic ones*" (Rawlins, 2015, p.219).

Table 3. Summary statistics (within and between variation)[*]

Variable		Mean	SD[†]	Min	Max	Obs[‡]
Business Unit level variables						
Growth	overall	-0.02	0.73	-9.26	12.21	53,258
(Δ ln sales)	between		0.42	-3.76	3.74	1,960
	within		0.71	-8.89	13.59	27.17
Size	overall	11.26	2.99	-1.46	18.63	56,070
(ln sales)	between		3.28	-0.21	18.25	2,090
	within		1.20	-0.47	17.75	26.83
Products	overall	2.19	2.38	1.00	31.00	56,070
	between		1.97	1.00	28.80	2,090
	within		0.56	-5.43	10.57	26.83
Variety	overall	2.49	4.75	0.00	62.00	56,070
	between		3.91	0.00	54.35	2,090
	within		1.26	-19.26	26.74	26.83
Parallel imports	overall	0.32	0.47	0.00	1.00	56,070
(sales proportion)	between		0.38	0.00	1.00	2,090
	within		0.22	-0.65	1.30	26.83
Generics	overall	0.23	0.42	0.00	1.00	56,070
(sales proportion)	between		0.40	0.00	1.00	2,090
	within		0.10	-0.74	1.20	26.83
Price	overall	-0.73	2.07	-7.28	7.98	56,070
(ln BU avg price)	between		2.16	-7.18	7.69	2,090
	within		0.37	-5.37	3.15	26.83
Class level (ATC4) variables						
Products	overall	9.03	13.30	1.00	131.00	13,717
	between		12.68	1.00	117.25	385
	within		1.86	-10.35	30.65	35.63
Variety	overall	10.15	21.66	0.00	220.00	13,717
	between		20.34	0.00	175.03	385
	within		3.96	-45.25	59.67	35.63
Herfindahl	overall	0.69	0.28	0.13	1.00	13,717
Index (HHI)	between		0.26	0.16	1.00	385
	within		0.10	0.19	1.23	35.63
Number	overall	4.23	3.96	1.00	29.00	13,717
of Firms	between		3.80	1.00	25.88	385
	within		0.74	-0.50	8.58	35.63

[*]Variation over time for a given individual is called *within (W)* variation, and variation across individuals (cross-section) is called *between (B)*. Statistics for *within* are computed by transforming the data by subtracting out the group mean and adding back in the overall mean.

[†]The overall variation can be approximately expressed as $s_O^2 \simeq s_W^2 + s_B^2$, where each component is computed as follows: $s_W^2 = \frac{1}{\sum_i T_i} \sum_i \sum_t (x_{it} - \overline{x}_i)^2$; $s_B^2 = \frac{1}{M-1} \sum_i (\overline{x}_i - \overline{x})^2$; $s_O^2 = \frac{1}{\sum_i T_i} \sum_i \sum_t (x_{it} - \overline{x})^2$.

[‡]Observations are listed as overall (M), over number of BUs (m) for between, and average number over time per BU ($\overline{T}$) for within.

and a cross-sectional SD of 0.10, the UK pharmaceutical industry for the prescription drug market appears fairly concentrated at the 4-digit ATC level. Nonetheless, about 32% of business units face some form of competition from parallel importers for one of their own products, and about 23% of BUs offer some generic product. The (log of) share weighted average price per unit of a drug (for

instance a tablet) across all drugs within a BU is -0.72, with greater variation between than within BUs (shares based on revenues).

5.2. **OLS estimates.** Table 4 provides estimates of select coefficients and marginal effects under alternative specifications of equation (3) while the full set of regression coefficients are given in Table B-1 in the Online Appendix.[15] Column (1) lists OLS results from a parsimonious specification that does not include lags of either the dependent variable (growth) or of products and varieties. However, this specification includes lagged size which, in contradiction to Gibrat's law of proportionate growth has a negative and significant coefficient indicating larger growth for smaller business units (see Simon and Bonini, 1958). This result is also in contradiction to the positive relation between size and productivity growth at the industry level reported in Pagano and Schivardi (2003).

Regarding the earlier stated hypotheses, note that the number of products and varieties have positive and significant coefficients which supports hypothesis one. The interaction term between products and varieties is negative and significant (albeit two orders of magnitude smaller) and provides support for hypothesis three that the two strategies are substitutes. Further, the coefficient on products is roughly twice as large as that on varieties and is consistent with hypothesis two. However, because the interaction term is negative and significant, we also provide marginal effects in the lower part of the table to see how large is the effect: a one unit increase in products is associated with a 26.4% increase in growth rate, while a one unit increase in varieties is associated with 14.3% increase in growth for the BU.[16] As we move through other columns in the table with more controls or alternative estimation strategies, the sign of the coefficients on products, varieties, the interaction term, and the relative magnitudes of the marginal effects listed at the bottom of the table help us check if hypotheses one-three still hold up or not. The test for the last hypothesis, that the marginal effect of additional products and varieties is larger for small BUs relative to large BUs is post-postponed till we re-estimate the model by sub-samples as given in the last 4 columns of the table.

In column (2) we provide OLS estimates of a distributed lagged model of the first difference form, which includes up to four lags of products, varieties, and the interaction of these terms (coefficients on lagged values are given in Table B-1). By adding these lagged variables to the model, the number

---

[15]All specifications include dummy variables for year and quarter, if the business unit has any sales due to generic products or due to parallel imports, and class level variables that include one period lag of number of products and varieties by competitors, number of competitors, and of HHI index for the class. These coefficients have been suppressed in the text but are given in the appendix.

[16]For the static model (column (1)), the marginal effect for product is $\partial g_{bt}/\partial p_{bt} = \beta_{10} + \beta_{30}v_{bt}$. We evaluate it at the sample average value of v_{bt} given in Table 3. The marginal effect with respect to variety is computed in a similar manner. The standard errors are computed via the 'delta' method.

Table 4. First difference growth models[†]

	OLS			IV			(4) By sub-samples (IV)			
	(1)	(2)	(3)	(4)	(5)	(6)	(4S)	(4M)	(4L)	(4X)
$s_{b,t-1}$: Lagged size	-1.06^a (0.016)	-1.05^a (0.016)	-1.05^a (0.034)	-0.76^a (0.099)	-0.71^a (0.12)	-0.71^a (0.12)	-0.77^a (0.13)	-0.77^a (0.18)	-1.12^a (0.35)	-0.79^a (0.10)
p_{bt}: Products	0.28^a (0.041)	0.30^a (0.044)	0.27^a (0.043)	0.26^a (0.044)	0.25^a (0.070)	0.25^a (0.070)	0.58^a (0.10)	0.40^a (0.12)	0.15^b (0.057)	0.15^a (0.024)
v_{bt}: Varieties	0.16^a (0.018)	0.15^a (0.018)	0.14^a (0.019)	0.14^a (0.019)	0.17^a (0.034)	0.18^a (0.035)	0.32^a (0.053)	0.25^a (0.055)	0.079^a (0.021)	0.074^a (0.012)
$p_{bt}v_{bt}$: Interaction ($\times 10^{-2}$) (Products×Varieties)	-0.70^a (0.13)	-0.70^a (0.13)	-0.65^a (0.12)	-0.63^a (0.12)	-0.77^a (0.26)	-0.80^a (0.26)	-3.40^a (0.90)	-1.60^a (0.54)	-0.30^a (0.11)	-0.31^a (0.064)
$g_{b,t-1}$: Growth (Lag 1)			-0.003 (0.026)	-0.28^a (0.098)	-0.34^a (0.11)	-0.34^a (0.11)	-0.36^a (0.12)	-0.19 (0.17)	0.22 (0.34)	-0.29^a (0.098)
x_{3bt}: log Price (Lag 2 log price)	0.0072 (0.027)	-0.0016 (0.024)	0.050^b (0.024)	0.027 (0.024)	0.017 (0.034)	0.020 (0.033)	-0.026 (0.028)	0.12^c (0.068)	0.052 (0.032)	0.040^c (0.023)
Marginal effects of products and varieties [‡‡]										
Products (Short run)		0.280^a (0.041)	0.201^a (0.032)	0.126^a (0.021)	0.104^a (0.029)	0.106^a (0.029)	0.246^a (0.043)	0.259^a (0.080)	0.130^b (0.052)	0.067^a (0.011)
Products (Long run)	0.264^a (0.039)	0.435^a (0.056)	0.309^a (0.044)	0.163^a (0.029)	0.146^a (0.049)	0.134^a (0.039)	0.357^a (0.080)	0.403^a (0.103)	0.186^a (0.067)	0.084^a (0.021)
Varieties (Short run)		0.130^a (0.016)	0.102^a (0.013)	0.063^a (0.009)	0.071^a (0.013)	0.071^a (0.014)	0.121^a (0.021)	0.152^a (0.034)	0.069^a (0.018)	0.033^a (0.005)
Varieties (Long run)	0.143^a (0.016)	0.188^a (0.022)	0.136^a (0.017)	0.065^a (0.013)	0.080^a (0.022)	0.072^a (0.017)	0.129^a (0.039)	0.209^a (0.048)	0.099^a (0.031)	0.031^a (0.008)
Observations	48,799	44,856	42,994	42,994	23,035	23,035	12,567	14,815	15,612	38,345
R-squared	0.510	0.507	0.508	0.502	0.512	0.512	0.556	0.462	0.425	0.561

[†]Standard errors are in parenthesis and clustered by business unit (firm-ATC4 combination). Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively. All regressions include additional controls at the class and BU level, as well as indicator variables for year and quarter (see Table B-1 in the appendix for detailed results).

Specification (1) does not include any lags, (2) is a distributed lag model and includes up to four lags of products, varieties and interactions. Specification (3) is an augmented distributed lag model that also includes four lags of the dependent variable (i.e., growth). Specification (4) treats only the first lag of growth as endogenous, (5) treats the first lag of growth, as well as products, and varieties, and their interaction and all four lags of each of these as endogenous variables (total 16), and specification (6) treats the first lag of growth and only the contemporaneous values of products, varieties and interaction as endogenous variables (total of 4 endogenous variables). Specifications (4S), (4M), and (4L) are same as (4), but on sub-samples by initial size of business unit being (s)mall, (m)edium and (l)arge, while (4X) restricts (4) to BUs that do not exit the sample.

[‡‡]Marginal effects are computed at the sample mean. They account for the interaction terms, the lagged values of the variable, as well as the lagged values of growth. Standard errors are computed using the delta method.

of observations drops from $\sim 49K$ to $\sim 45K$. Column (3) also provides OLS estimates on the first difference equation, but now extends the specification to an augmented distributed lagged model that includes four lags of the dependent variable (growth), and where once again, the coefficients on additional lags are given in the appendix. The number of observations further decreases to $\sim 43K$ due to the addition of these lagged variables to the model.

Note that in moving from a static specification in column (1) to the distributed lagged model in (2) or to a fully dynamic specification in (3), the coefficient on the size of BU does not change by much (-1.06 or -1.05 and significant), nor do the coefficients on contemporaneous values of products, varieties and the interaction terms. Marginal effects are given in the lower part of the table, but due to the lagged values of the variables in the dynamic growth models, it is possible to estimate long run and short run marginal effects separately.[17] Observe that the long run marginal effects in specification (3) are similar to those in the static model, i.e., 30.9% and 13.6% for products and varieties respectively, but the short run boost in revenue from an additional product or variety is 20.1% and 10.2%, respectively.[18] Importantly, hypotheses one-three still hold up with small changes in the coefficients of products, varieties and the interaction term, while the marginal effects of products are still larger than that of varieties both in the long and the short run.

5.3. **IV estimates.**

The next three columns re-estimate the dynamic growth model of (3) under alternative assumptions about the correlation of the error term with right hand side variables. Column (4) shows the results when only the first lagged value of growth is treated as endogenous, and accordingly instrumented for, but all other variables are assumed to be exogenous. The lagged value of growth is treated as endogenous because, as argued earlier, in the first difference form the first lagged value of the dependent variable becomes correlated with the error term by construction. Next, column (5) additionally treats the past and contemporaneous values of products, varieties and their interactions as endogenous, so there are 16 total endogenous variables, while column (6) treats only the first lagged value of growth and the contemporaneous values of products, varieties and interaction as correlated with the error term and hence only four variables are treated as endogenous in this model. Further, in columns (5) and (6) we restrict the sample to where a BU operates in more than one ATC4 class within the ATC2 due to availability of instruments (recall we construct the instruments for products and varieties for BU as the average value in other classes

---

[17]Column (2) is a distributed lag model and hence the long run marginal effect with respect to a product is given by $\partial g_{bt}/\partial p_{bt} = \beta_{10} + \beta_{30}v_{bt} + \sum_{j=1}^{4}(\beta_{1j} + \beta_{3j}v_{b(t-j)})$ while the short run marginal effect is the same expression as the one given for the static model in column (1). Marginal effects with respect to variety are computed in a similar manner. Column (3) is an augmented distributed lag (ADL) model and long run marginal effect with respect to product is $\partial g_{bt}/\partial p_{bt} = (\beta_{10} + \beta_{30}v_{bt} + \sum_{j=1}^{4}(\beta_{1j} + \beta_{3j}v_{b(t-j)}))/(1 - \sum_{m=1}^{4}\gamma_m)$ with an equivalent expression of marginal with respect to variety. In the case of the ADL models, if γ_m is not statistically significant, we set its value equal to zero in the computation of the marginal effect.

[18]For the static model in column (1), there is no distinction between short run and long run marginal effects since it is not built into the specification. However, when comparing the results between static and dynamic models (where the latter does have this distinction), we choose to interpret the results from the static model as those corresponding to the long run because, (i) the static model can be seen as an adjustment to long run equilibrium, and (ii) because the static model in (1) can itself be viewed as restricted version of (3), where the additional coefficients on the lags are constrained to be zero and hence the expression for the long run marginal effects in (2) and (3) correspond to the marginal effects in column (1).

and over time for the same firm, and hence the value is not available if the firm operates in only one ATC4 class).[19]

In terms of the main results, the largest change in estimated parameters and marginal effects comes when we move from column (3), where all variables are treated as exogenous, to column (4) which treats the first lag of the dependent variable as an endogenous variable. The coefficient on the lagged size drops from -1.05 to -0.76 (though still statistically significant) and that on lagged growth increases in magnitude from -0.003 and not significant to -0.28 and significant and with almost no change in the lagged coefficients for products and varieties. However the long run marginal effects with respect to products and varieties (computed at the sample mean), drop to roughly half their previous values, i.e., from 30.9% and 13.6% in column (3) to 16.3% and 6.5% in column (4).

The results in the next two columns with either all variables related to products and varieties and their interactions are endogenous (column 5), or just their contemporaneous values are endogenous (column 6) are generally similar, while the marginal effects with respect to products decrease slightly and those with respect to varieties increase. We verified that this additional change in coefficients and marginal effects computed at the mean of the sample is mostly driven by change in the sample (instead of nearly $\sim 43K$ observations in (4), we have about $\sim 23K$ in (5) and (6)) rather than due to any additional endogeneity of products and packs by restricting (4) to the sample.[20] Regardless of which specification we use, coefficient signs and marginal effects are consistent with hypotheses one-three (we look at difference in magnitude of marginal effects with respect to products and varieties, and if the difference is statistically significant, in the next section).

For specifications (4), (5) and (6), results related to the first stage F-statistics (weak instruments test), under identification and over identification are given in the Online Appendix in Table B-2. In all cases the first stage F-statistics suggest that the instruments are not weak as the relevant F-values are always above 10, and in all three specifications, the null of under identification is rejected and in models (5) and (6) the null of over identification in not rejected (model 4 is just identified so the test is not available). Similar tests for the specifications by sub-samples (4S) through (4X) are also summarized in the same table and do not point to any problems with these instruments.[21]

---

[19]An alternative set of instruments discussed earlier is to include BUs that operate within only a single class and to construct instruments using the average value from other time periods. Results from these alternative instruments are qualitatively similar, but since their validity is suspect, we prefer to restrict the sample to firms that operate in more than one ATC4 class, and not use variation over time alone to construct the instruments.

[20]For instance, if we re-estimate (4) with the restricted sample, the four marginal effects for specification (4) become 0.104, 0.135, 0.054, and 0.055 respectively, which are very close to those of specification (6).

[21]Since our additional instruments for specifications (5) and (6) rely on inter-class, intra-firm correlation in capabilities, clustering the standard errors at sub-firm level may be masking a weak instruments problem (we cluster at the BU level which is defined via Firm*ATC4 interaction). To check if this is so, we re-estimated all the models but clustered either at the firm level or at the therapeutic class level. Neither of these alternative clusterings resulted in first stage statistics that
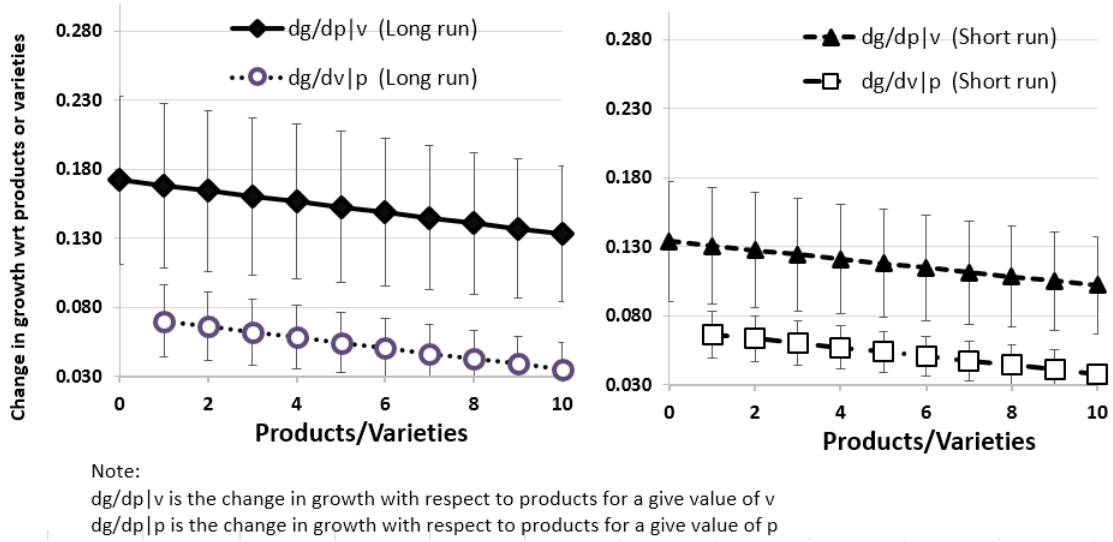
Observe also that coefficients related to products and pack varieties do not change much once we treat these additional variables as endogenous. Indeed a Hausman test between specifications (4) and (5) (i.e. the vector of contrasts between these specifications and the so-called endogeneity test) does not reject the null of the exogeneity of these additional 15 variables (the $\chi^2(15)$ is 17.22 with an associated p-value of 0.306). This does not mean that the endogeneity is not present, but rather if present, it is mild enough that it does not appear to have a significant effect on the estimates of the regression coefficients. Given that there is not much difference in the estimated parameters and marginal effects in the three models (especially after we account for sample size differences), we use specification (4) as the preferred model and discuss further results in light of this specification (we do however also provide additional results by specification (6) in the appendix).

5.4. **Marginal effects.** Due to the interaction term between products and varieties (which are significant in all models), the marginal effects are not constant. For instance, the marginal effect with respect to products is a function of varieties, and similarly the other way around, and this is true for both the long run and short run marginal effects. Figure 3 plots the estimated marginal effect of products for a range of observed values of varieties, and similarly, the marginal effect of varieties for a range of values of products. To be clear, the graph shows $\partial g_{bt}/\partial p_{bt}$ on vertical axis plotted against values of v_{bt} and also $\partial g_{bt}/\partial v_{bt}$ plotted against p_{bt} (also recall that our pack variety measure starts at zero as it is defined as total number of products minus total number of packs, and hence for a BU with one pack for each product, pack variety is zero). The marginal effects are for both the long run (left panel) and short run (right panel), and the vertical error bars show the 95% confidence intervals.

Consistent with hypothesis one, each of the four marginal effects are positive and significantly different from zero for a range of observed values of products and varieties and not just at the sample mean values of 2.19 and 2.49. Equally important, the negative slopes of these graphs imply that the impact on growth from an additional product is smaller if the BU has many pack varieties than if it had fewer pack varieties (hypothesis three). Similarly, the marginal with respect to pack varieties is diminishing in products. Both in the long and shot run, the marginal effects with respect to products is higher than that of varieties and the non-overlapping error bars indicate that these differences in marginal effects are statistically significant over a range of values, i.e., we do not reject hypothesis two. However, the figure also shows that while there is a (statistical) difference in the long and short run marginal effect of additional products, there is virtually none in the long and

---

would indicate that the instruments are weak. The entire first stage results, along with under and over identification tests with alternative clustering at either the firm level or the therapeutic class level are summarized in Table B-3. Overall, the statistical tests summarized in these tables lend support to these instruments as valid and relevant, and that in first difference form (and after treating lagged growth as endogenous) there is perhaps no additional endogeneity concern for products and packs.

Figure 3. Marginal Effects – Products and Varieties

Note:
dg/dp|v is the change in growth with respect to products for a give value of v
dg/dp|p is the change in growth with respect to products for a give value of p

short run for varieties. In turn, these imply that the boost in sales from additional pack varieties are smaller but more immediate, whereas those from introducing additional products are larger, but also over a longer period of time.

5.5. **By subsamples.** We next turn to hypothesis four. So far our analysis controls for size of the business unit but does not allow for the impact of products and varieties on growth to vary by the size of the business unit (see Belenzon and Patacconi (2014) for differential effects of patents and publications by firm size). To study this effect, one possible extension is to allow for an interaction between (lagged) size and products and varieties so that the marginal effects can vary by size. Instead, due to the richness of our data, we allow for more flexible version by re-estimating specification (4) on separate sub-samples by size, which in turn let all coefficients vary by size, rather than just those related to products and varieties. Columns (4S), (4M) and (4L) in Table 4 display estimates from sub-samples by business unit size classified as small, medium and large respectively (recall that the classification is based on tertile distribution of the initial log size). Appendix (A.4) and Table A-4 provide additional results for sub-samples based on specification (6).

The change in the estimated coefficients for the relevant variables are substantial, both when the comparison is done within the sub-sample of class sizes and when it is set against the entire sample displayed in column (4). The coefficients on products, varieties and interactions increase in magnitude for small business units and decrease for large business units. For instance, in column (4) the coefficient on varieties is 0.14 but this increases to 0.32 and 0.25 for small and medium size BUs respectively (4S, 4M), and decreases to 0.08 for large BUs (4L). Similarly, the coefficient on products also increases for small and medium size BUs (0.58 and 0.40 respectively) but decreases for

the large BUs (0.15). In fact, consistent with hypothesis four, the long run marginal effect of new products is positive and much greater for small BUs than for large BUs. On the other hand, the marginal impact of additional variety on growth is much smaller and was not statistically different for small vs large BUs, and hence does not support the hypothesis.

While several studies have found evidence supporting growth by small firms (see for instance Calvo, 2006) or at least for short periods after birth (e.g. Lotti et al., 2003), our results shed further insight into this mechanism for growth: the returns in terms of growth from introducing a new product are much larger for small BUs than for larger ones. This could be because smaller firms (in our case smaller BUs) introduce more radical innovations relative to their counterparts. From our data we cannot attest to the type of additional drugs launched by BUs, i.e., how radically differentiated they are in comparison to their existing portfolio of drugs. However, we find that smaller BUs focus more on new drugs relative to pack varieties, while larger BUs have greater pack variation per product (packs per product ratio is 2.68 for large BUs and 1.70 for small BUs). Additionally, we find that when a small BU introduces a new drug, there are relatively fewer pre-existing drugs in the ATC class, relative to when a large BU introduces a new drug: when a small BU introduces a new drug, the median (mean) number of drugs already in class are 13.5 (22.6), while similar numbers for a large BU are 28 (38.5). If relative entry sequence is indicative of novelty, these statistics are suggestive that small BUs are willing to take risks and hence perhaps introduce more radically different drugs. This would be consistent with Akcigit (2010) and Akcigit and Kerr (2018) who find that smaller firms produce higher quality innovations (as measured by citation counts), or more 'external' vs 'internal' innovations (i.e., new products vs improvements of product lines) and with Caggese (2019) who reports that small firms undertake more radical innovation.

Finally, in column (4X) we consider one additional sub-sample; only those business units that survive all the way to the end of our observational window. Business units that do not survive are smaller, have fewer number of products and packs, and experience negative growth compared to their surviving counterparts (and these differences are statistically significant), which is similar to findings reported elsewhere in the literature (see for instance Agarwal, 1997). Accordingly, we find that the coefficients on products and pack varieties, as well as marginal effects in the short and long run are similar to those reported earlier for large business units. This result is consistent with that reported by Disney et al. (2003) who find that in the UK manufacturing sector, small entrants are more likely to fail than their larger counterparts, and that the hazard declines if there is fast growth. Our results using BU analysis suggest that this fast growth for small firms may be precisely due to product and pack launches as shown in column (4S).

In line with hypothesis one, we find that both types of innovations have a positive impact on sales growth. Furthermore, the findings that a new product leads to substantial growth in the short and

long run, whereas a new pack variety only produces a short-term effect with no additional impact on growth over the long run confirm hypothesis two. The negative coefficient on the interaction between products and varieties corroborates hypothesis three. We also find that in the British pharmaceutical sector, smaller business units grew more than larger ones. Equally important, as put forward in hypothesis four, for an existing combination of products and varieties, an additional product may be more profitable, vis-a-vis opportunity for growth, for the smaller business units than for larger ones. On the other hand, the difference in marginal effects with respect to pack variety is much smaller relative to products, and also does not differ by size of the business unit.

5.6. **Quantity and price.** The foregoing analysis shows that product and pack varieties are positively associated with revenue sales growth. Since revenue is quantity times price, to check if this association in growth is via increase in price, in quantity demanded or both, we estimated auxiliary regressions using price and quantity indexes on counts of products and packs. The quantity index was computed as the ratio of revenue sales to a price index for the BU, where the latter itself was computed as a share weighted average price per unit. Simple reduced form regressions (not shown here in interest of space) of log of quantity index and log of price index on products, pack varieties, and time trends show strong positive associations of these variables on both the price and quantity (regressions included BU fixed effects, polynomial time trends and clustered standard errors). Further, we also estimated a growth model identical to the one in specifications (4) and (6) above, but using change in quantity index rather than revenue as a measure of growth. For the equivalent of specification (4), estimates for quantity growth model were very similar to the revenue growth model, with marginal effects being 9.0% and 14.7% for products (short run and long run respectively) and 5.5% and 6.2% for varieties (short run and long run respectively), indicating that new introductions affect quantity demanded as well as prices.[22]

5.7. **Robustness to OTC and non-reimbursable drugs.** Recall that we have excluded all drugs from analysis that were over-the-counter or not reimbursable by NHS even though sales for those are available in our data. Primary reason for exclusions was that IMS does not cover direct sales from manufacturers to hospitals or to non-pharmacy outlets and hence our data on OTC and non-reimbursable drugs is incomplete. Habl et al. (2006, p.707) report that in 2005 total OTC or self-medication market represented about 12.03% of all pharmaceutical sales in the UK. For the same year, our data which only records sales to pharmacies and doctors' offices, shows that OTC and non-reimbursable make up 7.85% of all sales (for the full 10 years they are 13.63% of all sales in our data). Thus, based on 2005 sample, our data would be missing about 4.2% of sales which are presumably to hospitals and grocery stores etc. Additionally, we also omitted these sales as

---

[22]The marginal effects for equivalent of specification (6) for quantity growth were 8.5% and 14.2% for products (short run and long run respectively), and 7.1% and 8.0% for varieties (short run and long run respectively).

these outlet markets may be fundamentally different from the prescription drug market covered by pharmacies. Nonetheless, ignoring these outside sales not recorded in the data, we re-estimated the models by including OTC and non-reimbursable drugs. Marginal effects from specifications (1), (4) and (6) are given in Table 5 (and for ease of comparison, those from the original data are also listed in the same table). The marginal effects are fairly close to the original estimates though slightly smaller than before.

TABLE 5. Comparison with OTC/Non-Reimbursable Included*

	(1)		(4)		(6)	
	Original	With	Original	With	Original	With
Products (Short run)			0.126^a (0.021)	0.108^a (0.015)	0.106^a (0.029)	0.104^a (0.022)
Products (Long run)	0.264^a (0.039)	0.239^a (0.028)	0.162^a (0.029)	0.137^a (0.022)	0.134^a (0.039)	0.138^a (0.032)
Varieties (Short run)			0.063^a (0.009)	0.053^a (0.006)	0.071^a (0.014)	0.057^a (0.010)
Varieties (Long run)	0.143^a (0.016)	0.125^a (0.013)	0.065^a (0.013)	0.053^a (0.011)	0.072^a (0.017)	0.059^a (0.011)

* 'With' refers to sales data inclusive of all OTC and other non-reimbursable drugs sold in UK pharmacies and doctors' offices. 'Original' refers to the original data not inclusive of these additional sales.

5.8. **Robustness to generics and unidentified generic manufacturers.** In the preceding results, the regression coefficient for generic products is not significant (x_{2bt}). Note that this variable is coded as a 1/0 dummy indicating if the BU has any sales due to generic drugs. If instead we construct it as share of sales that are due to generic drugs, the regression coefficient changes but is still not significant in any of the models. Further, the marginal effects for products and varieties are virtually the same (we verified this for specifications (1), (4) and (6)). Similarly, if we drop this variable from the regressions entirely, the marginal effects reported earlier remain unchanged. In interest of space we do not show them here.

However, recall that our data does not separately identify all generic manufacturers (see footnote 8 in the main text): for some manufacturers producing only generic drugs we do not know their identity, and hence all such manufacturers within a given ATC4 class are treated as one firm. Clearly any additional drugs introduced by these BUs are not 'innovations', since by definitions they are generic drugs. Same goes for drug introductions by manufacturers with known identify when they introduce a drug which is a generic. If we remove all such observations and re-estimate the models, the marginal effects for products increase slightly, but not much for pack varieties. We

TABLE 6. Comparison with generics and unknown manufacturers excluded[*]

	(1)		(4)		(6)	
	Original	Without	Original	Without	Original	Without
Products (Short run)			0.126^a	0.153^a	0.106^a	0.129^c
			(0.021)	(0.036)	(0.029)	(0.067)
Products (Long run)	0.264^a	0.338^a	0.162^a	0.212^a	0.134^a	0.162^b
	(0.039)	(0.067)	(0.029)	(0.051)	(0.039)	(0.080)
Varieties (Short run)			0.063^a	0.073^a	0.071^a	0.082^a
			(0.009)	(0.013)	(0.014)	(0.020)
Varieties (Long run)	0.143^a	0.160^a	0.065^a	0.099^a	0.072^a	0.106^a
	(0.016)	(0.023)	(0.013)	(0.020)	(0.017)	(0.028)

[*] 'Without' refers to sales data not including unknown generic BUs or observations from BUs if they produce any generic drugs. 'Original' refers to the original data not inclusive of these additional sales.

show in Table 6 the marginals from specifications (1), (4) and (6) with and without these additional observations.

6. CONCLUSIONS

Our paper contributes to the innovation-growth literature in various ways. First, we map innovative output rather than input to growth, and highlight the fact that in the pharmaceutical sector we have different types of innovations which differ in their appropriability conditions, and hence may have differential impact on growth (Tavassoli and Karlsson, 2015). Product innovations are drugs with new molecules or formulations, which often enjoy patent and marketing exclusivity, and can serve a patient base for which earlier drugs were less suitable. These types of innovations differ from marketing innovations, such as a new pack that varies by size or strength and could be introduced due to price discrimination motivation. These pack varieties will not necessarily have any additional marketing exclusivity associated with them. Heterogeneity in these sources of innovations can have an asymmetric impact on revenue growth, and in this work, we emphasize this aspect.

A second contribution is that we refine the empirical analysis at the business unit level rather than at firm level. This distinction is important, since the contribution of each innovation to a business unit growth can get lost at the firm level as many business units not innovating can wash it out. This in fact may be a reason why the literature has not found a robust link between innovative activities and firm performance (for a review, see Coad, 2009, Audretsch et al., 2014). Aggregation across heterogenous sub-markets in which a pharmaceutical firm may be operating, i.e., vastly different therapeutic classes, may mask the true relationship between innovative outputs and growth, a point

made earlier by Corsino and Gabriele (2011) in a different setting. A fundamental question then is whether the lessons learnt in this current research about the link between innovation and growth at the business unit level imply more generally similar relationship at the firm level? Our data indicate that large business units are associated to firms that are almost double the size of firms associated to small business units. Industries where sub-markets have little or no heterogeneity in products or services, firm and business unit level analysis ought to give similar conclusions. However, when there is significant variation in product sub-markets, inferring links between innovation and growth from business unit level analysis to firm level may not be straight forward. Nonetheless, in our view, in such cases the appropriate analysis is at a more disaggregated level. For instance, in telecommunications, it may be more useful to study the output innovations within sub-sectors, such as wireless communications, processing systems, long distance carriers, broadband and data services etc. and linking those to specific units of AT&T, Verizon, Vodafone etc. that are active in these areas, rather than overall innovative outputs and firm level revenue changes which could be driven by many other factors in these firms.

Finally, our analysis adds to the literature on growth-innovation link by firm size, albeit we do so in the context of pharmaceutical business units and by type of innovations (see Coad, 2009). As pointed out in Pagano and Schivardi (2003), small firms are important for job creation and growth, as they intensify competition. In our exposition, we showed that small entities (small business units) are able to get higher returns from innovations, vis-à-vis growth opportunities, particularly product innovations, compared to their larger counter parts. In part this may be because they introduce more innovative products, as for instance indicated by entry order of their products in the therapy class (but it could also be due to other factors not analyzed here, for instance better management or market intelligence). The fact that small business units receive a higher return from product innovations than large business units are evocative of a dynamic market. Also, our finding that new drug introductions generate long run business-unit revenue growth is suggestive that the pharmaceutical market in the UK, through appropriability conditions, is endowed with powerful tools to expand, even when patient bases in therapy class may be slow to grow.

After an initial descriptive analysis, we used reduced form linear dynamic revenue growth models to explore how the number of drugs and pack varieties trigger business units' growth. Looking at both long and short run, we find that both product and marketing innovations have a significant impact on revenue growth, with a magnitude larger for new drugs than new packs. A finding of interest is that new products generate persistent growth – evidence that product innovation is a robust driving force for business unit growth – whereas marketing innovations contribute only to short run growth. However, as we emphasize in the paper, producing only short run growth is not *per se* a bad strategy. In fact, due to price regulation for branded drugs, marketing innovations can

be a profitable short-term business strategy, as the cost of introducing a new pack is likely not to exceed that of introducing a new drug.

There are several limitations of our work that suggest extensions. While we control for competition using a HHI index at the class level, we have not allowed for an interaction between heterogeneous innovation and competition. Mazzucato and Parris (2015) have studied (for high-growth firms) the effect of R&D investments on growth during periods of intense and soft competition. One could go one-step further and test whether the heterogeneous effect of different innovative outputs on growth is affected by competition, or market structure more generally. This could be executed using different periods of competition, as in Mazzucato and Parris (2015), or by interacting an index of competition with innovative outputs/inputs. Alternatively, it would be useful to study the growth-innovation link in the context of endogenous market power and exit of non-innovating business units. A related possible extension of this study is an investigation of the asymmetric effect of innovation for high-growth and low-growth business units, using a quantile regression as in Coad and Rao (2008) and Capasso et al. (2015). This experiment could be extended to compare growth for top firms/business units (top 10, top 20, and top 50) and study whether skewness of size is informative on the impact of innovative outputs on growth.

In our analysis we focused on innovation affecting growth, and while we acknowledge the reverse causality and attempt to correct for this source of endogeneity, we do not explicitly model what type of business units are more likely to introduce innovations. Nor do we fully investigate whether the novelty of innovations differ by business unit characteristics, or how these characteristics may interact with market structure to determine future innovation. A conclusion from this study is to facilitate introduction of new drugs relative to pack variation. However, there is need for further research that separates out the types of new drug introduction, distinguishing between radical innovations and small incremental variations of existing drugs such as me-too drugs, or those introduced as part of a product hopping strategy (Hemphill and Sampat, 2012, Bokhari and Fournier, 2013). This latter strategy may also have some anticompetitive effects and it would be important to test for the separate effects that the two have on growth, with particular focus on small versus large firms (or business units). We hope to investigate some of these issues in future research.

References

Acs, Zoltan J. and David B. Audretsch, "Innovation in large and small firms: an empirical analysis," *The American Economic Review*, September 1988, *78* (4), 678–690.

_____ **and** _____ , *Innovation and Small Firms*, Cambridge, MA: MIT Press, 1990. ISBN: 9780262011136.

Agarwal, Rajshree, "Survival of firms over the product life cycle," *Southern Economic Journal*, January 1997, *63* (3), 571–584.

Aghion, Philippe and Peter Howitt, "A model of growth through creative destruction," *Econometrica*, March 1992, *60* (2), 323–351.

Akcigit, Ufuk, "Firm size, innovation dynamics and growth," Working paper (University of Pennsylvania) 2010.

_____ **and William R. Kerr**, "Growth through heterogeneous innovations," *Journal of Political Economy*, 2018, *126* (4), 1374–1443.

Anderson, Theodore W. and Cheng Hsiao, "Formulation and estimation of dynamic models using panel data," *Journal of Econometrics*, January 1982, *18* (1), 47–82.

Arcidiacono, Peter, Paul Ellickson, Peter Landry, and David B. Ridley, "Pharmaceutical followers," *International Journal of Industrial Organization*, September 2013, *31* (5), 538–553.

Audretsch, David B., Alex Coad, and Agustí Segarra, "Firm growth and innovation," *Small Business Economics*, December 2014, *43* (4), 743–749.

_____ **and Talat Mahmood**, "The rate of hazard confronting new firms and plants in U.S. manufacturing," *Review of Indistrual Organization*, February 1994, *9* (1), 41–56.

Baldwin, John R. and Paul K. Gorecki, "Firm entry and exit in the Canadian manufacturing sector, 1970-1982," *The Canadian Journal of Economics*, May 1991, *24* (2), 300–323.

Banbury, Catherine M. and Will Mitchell, "The effect of introducing important incremental innovations on market share and business survival," *Strategic Management Journal*, Summer 1995, *16* (S1), 161–182.

Becheikh, Nizar, Réjean Landry, and Nabil Amara, "Lessons from innovation empirical studies in the manufacturing sector: A systematic review of the literature from 1993–2003," *Technovation*, May-June 2006, *26* (5-6), 644–664.

Belenzon, Sharon and Andrea Patacconi, "How does firm size moderate firms' ability to benefit from invention? Evidence from patents and scientific publications," *European Management Review*, February 2014, *11* (1), 21–45.

Bernard, Andrew B., Emily J. Blanchard, Ilke van Beveren, and Hylke Vandenbussche, "Carry-Along Trade," *The Review of Economic Studies*, March 2018, *86* (2), 526–563.

Bokhari, Farasat A. S. and Gary M. Fournier, "Entry in the ADHD drugs market: Welfare impact of generics and me-toos," *Journal of Industrial Economics*, June 2013, *61* (2), 340–393.

Bottazzi, Giulio and Angelo Secchi, "Growth and diversification patterns of the worldwide pharmaceutical industry," *Review of Industrial Organization*, March 2005, *26* (2), 195–216.

_____, **Giovanni Dosi, Marco Lippi, Fabio Pammolli, and Massimo Riccaboni**, "Innovation and corporate growth in the evolution of the drug industry," *International Journal of Industrial Economics*, July 2001, *19* (7), 1161–1187.

Cabral, Luís M. B. and José Mata, "On the evolution of the firm size distribution: Facts and theory," *American Economic Review*, September 2003, *93* (4), 1075–1090.

Caggese, Andrea, "Financing constraints, radical versus incremental innovation, and aggregate productivity," *American Economic Journal: Macroeconomics*, 2019, *11* (2), 275–309.

Calvo, José L., "Testing Gibrat's Law for small, young and innovating firms," *Small Business Economic*, August 2006, *26* (2), 117–123.

Capasso, Marco, Tania Treibich, and Bart Verspagen, "The medium-term effect of R&D on firm growth," *Small Business Economics*, June 2015, *45* (1), 29–62.

Coad, Alex, *The growth of firms: A survey of theories and empirical evidence*, Northampton, Massachusetts, USA: Edward Elgar Publishing, 2009.

_____, **Agustí Segarra, and Mercedes Teruel**, "Innovation and firm growth: Does firm age play a role?," *Research Policy*, March 2016, *45* (2), 387–400.

_____ **and Rekha Rao**, "Innovation and firm growth in high-tech sectors: A quantile regression approach," *Research Policy*, May 2008, *37* (4), 633–648.

_____ **and** _____ , "Firm growth and R&D expenditure," *Economics of Innovation and New*

Technology, December 2010, *19* (2), 127–145.

Connelly, Dawn, "A breakdown of the over-the-counter medicines market in Britain in 2016," *The Pharmaceutical Journal*, April 2017, *298* (7900).

Conner, Kathleen Reavis, "Strategies for product cannibalism," *Strategic Management Journal*, Summer 1988, *9* (s1), 9–26.

Corsino, Marco and Roberto Gabriele, "Product innovation and firm growth: evidence from the integrated circuit industry," *Industrial and Corporate Change*, February 2011, *20* (1), 29–56.

Cucculelli, Marco and Barbara Ermini, "New product introduction and product tenure: What effects on firm growth?," *Research Policy*, June 2012, *41* (5), 808–821.

de Frutos, Maria-Angeles, Carmine Ornaghi, and Georges Siotis, "Competition in the pharmaceutical industry: How do quality differences shape advertising strategies?," *Journal of Health Economics*, January 2013, *32* (1), 268–285.

Del Monte, Alfredo and Erasmo Papagni, "R&D and the growth of firms: empirical analysis of a panel of Italian firms," *Research Policy*, June 2003, *32* (6), 1003–1014.

Demirel, Pelin and Mariana Mazzucato, "Innovation and firm growth: is R&D worth it?," *Industry and Innovation*, January 2012, *19* (1), 45–62.

Deschryvere, Matthias, "R&D, firm growth and the role of innovation persistence: an analysis of Finnish SMEs and large firms," *Small Business Economics*, Dececmber 2014, *43* (4), 767–785.

DiMasi, Joseph A., Henry G. Grabowski, and Ronald W. Hansen, "Innovation in the pharmaceutical industry: New estimates of R&D costs," *Journal of Health Economics*, May 2016, *47*, 20–33.

Disney, Richard, Jonathan Haskel, and Ylva Heden, "Entry, exit and establishment survival in UK manufacturing," *The Journal of Industrial Economics*, march 2003, *51* (1), 91–112.

Dunne, Paul and Alan Hughes, "Age, size, growth and survival: UK companies in the 1980s," *The Journal of Industrial Economics*, June 1994, *42* (2), 115–140.

Dunne, Timothy, Mark J. Roberts, and Larry Samuelson, "Patterns of firm entry and exit in U.S. manufacturing industries," *The RAND Journal of Economics*, winter 1988, *19* (4), 495–515.

Ellison, Glenn and Sara Fisher Ellison, "Strategic entry deterrence and the behavior of pharmaceutical incumbents prior to patent expiration," *American Economic Journal: Microeconomics*, February 2011, *3* (1), 1–36.

Evans, David S., "The relationship between firm growth, size, and age: Estimates for 100 manufacturing industries," *The Journal of Industrial Economics*, June 1987, *35* (4), 567–581.

———, "Tests of alternative theories of firm growth," *Journal of Political Economy*, August 1987, *95* (4), 657–674.

Flor, Maria Luisa and María José Oltra, "Identification of innovating firms through technological innovation indicators: an application to the Spanish ceramic tile industry," *Research Policy*, March 2004, *33* (2), 323–336.

Frank, Richard G. and David S. Salkever, "Pricing patent loss and the market for pharmaceuticals," *Southern Economic Journal*, October 1992, *59* (2), 165–179.

Freel, Mark S., "Do small innovating firms outperform non-innovators?," *Small Business Economics*, May 2000, *14* (3), 195–210.

Gambardella, Alfonso et al., *Science and innovation: The US pharmaceutical industry during the 1980s*, Cambridge University Press, 1995.

Geroski, Paul A., "Understanding the implications of empirical work on corporate growth rates," *Managerial and Decision Economics*, 2005, *26* (2), 129–138.

——— **and Mariana Mazzucato**, "Learning and the sources of corporate growth," *Industrial and Corporate Change*, August 2002, *11* (4), 623–644.

——— **and Saadet Toker**, "The turnover of market leaders in UK manufacturing industry,

1979-86," *International Journal of Industrial Organization*, 1996, *14*, 141–158.

_____ **and Stephen Machin**, "Do innovating firms outperform non-innovators?," *Business Strategy Review*, June 1992, *3* (2), 79–90.

_____ , _____ , **and Christopher F. Walters**, "Corporate growth and profitability," *The Journal of Industrial Economics*, June 1997, *45* (2), 171–189.

Grabowski, Henry, John Vernon, and Joseph A. DiMasi, "Returns on research and development for 1990s new drug introductions," *PharmacoEconomics*, December 2002, *20* (Supplement 3), 11–29.

Graves, Samuel B. and Nan S. Langowitz, "Innovative productivity and returns to scale in the pharmaceutical industry," *Strategic Management Journal*, November 1993, *14* (9), 593–605.

Habl, Claudia, Katja Antony, Danielle Arts, Michael Entleitner, Barbara Fröschl, Christine Leopold, Heidi Stürzlinger, Sabine Vogler, and Marion Weigl, *Surveying, assessing and analysing the pharmaceutical sector in the 25 member states*, ÖBIG Health Economics, Commissioned by European Union - DG Competition, 2006.

Hall, Bronwyn H., "The relationship between firm size and firm growth in the U.S. manufacturing sector," *Journal of Industrial Economics*, June 1987, *35* (4), 583–600.

Hemphill, C. Scott and Bhaven N. Sampat, "Evergreening, patent challenges, and effective market life in pharmaceuticals," *Journal of Health Economics*, March 2012, *31* (2), 327–339.

Henderson, Rebecca and Iain Cockburn, "Scale, scope, and spillovers: the determinants of research productivity in drug discovery," *Rand Journal of Economics*, Spring 1996, *27* (1), 32–59.

Hurwitz, Mark A. and Richard E. Caves, "Persuasion or information? Promotion and the shares of brand name and generic pharmaceuticals," *The Journal of Law and Economics*, October 1988, *31* (2), 299–320.

Huskamp, Haiden A., Julie M. Donohue, Catherine Koss, Ernst R. Berndt, and Richard G. Frank, "Generic Entry, Reformulations, and Promotion of SSRIs," *Pharmacoeconomics*, July 2008, *26* (7), 603–616.

Jaska, A., L. Westbrook, E. Rubinstein, K. Danniel, and Y.S. Ho, "NICE restrictiveness compared to the market authorization," *Value in Health*, November 2014, *17* (7), A438–A439.

_____ , _____ , _____ , _____ , **and** _____ , "NICE restrictiveness compared to the market authorization in oncology and non-oncology reviews," *Value in Health*, November 2014, *17* (7), A660–A661.

Jovanovic, Boyan, "Selection and the evolution of industry," *Econometrica*, May 1982, *50* (3), 649–670.

Kanavos, Panos, "Overview of pharmaceutical pricing and reimbursement regulation in Europe," *Japanese Pharmacology and Therapeutics*, 2003, *31* (10), 819–838.

Landsman, Vardit, Isabel Verniers, and Stefan Stremersch, "The successful launch and diffusion of new therapies," in Min Ding, Jehoshua Eliashberg, and Stefan Stremersch, eds., *Innovation and marketing in the pharmaceutical industry*, Vol. 20 of *International Series in Quantitative Marketing*, New York, NY: Springer, 2003, chapter 7, pp. 189–223.

Lotti, Francesca, Enrico Santarelli, and Marco Vivarelli, "Does Gibrat's Law hold among young, small firms?," *Journal of Evolutionary Economics*, August 2003, *13* (3), 213–235.

Manchanda, Puneet and Elisabeth Honka, "The effects and role of direct-to-physician marketing in the pharmaceutical industry: an integrative review," *Yale Journal of Health Policy, Law, and Ethics*, 2005, *5* (2), 785–822.

Matia, Kaushik, Dongfeng Fu, Sergey V. Buldyrev, Fabio Pammolli, Massimo Riccaboni, and H. Eugene Stanley, "Statistical properties of business firms structure and growth," *EPL (Europhysics Letters)*, 2004, *67* (3), 498.

Mazzucato, Mariana and Stuart Parris, "High-growth firms in changing competitive environments: the US pharmaceutical industry (1963 to 2002)," *Small Business Economics*, January

2015, *44* (1), 145–170.

NHS Digital, *Prescribing costs in hospitals and the community, England 2017/18* November 2018.

Orsenigo, Luigi, Fabio Pammolli, and Massimo Riccaboni, "Technological change and network dynamics: lessons from the pharmaceutical industry," *Research policy*, 2001, *30* (3), 485–508.

Pagano, Patrizio and Fabiano Schivardi, "Firm size distribution and growth," *The Scandinavian Journal of Economics*, June 2003, *105*, 255–274.

Paul, Steven M., Daniel S. Mytelka, Christopher T. Dunwiddie, Charles C. Persinger, Bernard H. Munos, Stacy R. Lindborg, and Aaron L. Schacht, "How to improve R&D productivity: the pharmaceutical industry's grand challenge," *Nature Reviews Drug Discovery*, 2010, *9* (3), 203.

Phillips, Bruce D. and Bruce A. Kirchhoff, "Formation, growth and survival; Small firm dynamics in the U.S. economy," *Small Business Economics*, March 1989, *1* (1), 65–74.

Rawlins, Michael D, "Generic prescribing: unfinished business," *The Lancet*, 2015, *385* (9964), 219.

Regan, Tracy L., "Generic entry, price competition, and market segmentation in the prescription drug market," *International Journal of Industrial Organization*, July 2008, *26* (4), 930–948.

Roper, Stephen, "Product innovation and small business growth: a comparison of the strategies of German, UK and Irish companies," *Small Business Economics*, 1997, *9* (6), 523–537.

Ruggeri, Kai and Ellen Nolte, *Pharmaceutical pricing: The use of external reference pricing*, Santa Monica, CA: RAND Corporation, 2013.

Scannell, Jack W., Alex Blanckley, Helen Boldon, and Brian Warrington, "Diagnosing the decline in pharmaceutical R&D efficiency," *Nature Reviews Drug discovery*, 2012, *11* (3), 191.

Scott Morton, Fiona and Margaret Kyle, "Markets for pharmaceutical products," in Mark V. Pauly, Thomas G. Mcguire, and Pedro P. Barros, eds., *Handbook of Health Economics*, Vol. 2, Amsterdam, North-Holland: Elsevier Science Inc., 2011, chapter 12, pp. 763–823.

Simon, Herbert A. and Charles Bonini, "The size distribution of business firms," *American Economic Review*, September 1958, *48* (4), 607–617.

Stam, Erik and Karl Wennberg, "The roles of R&D in new firm growth," *Small Business Economics*, June 2009, *33* (1), 77–89.

Tavassoli, Sam and Charlie Karlsson, "Persistence of various types of innovation analyzed and explained," *Research Policy*, month 2015, *44* (10), 1887–1901.

Thomke, Stefan and Walter Kuemmerle, "Asset accumulation, interdependence and technological change: evidence from pharmaceutical drug discovery," *Strategic Management Journal*, July 2002, *23* (7), 619–635.

Torjesen, Ingrid, "Drug development: the journey of a medicine from lab to shelf," *The Pharmaceutical Journal*, May 2015. Online, URI: 20068196.

Ward, D.J., A. Shiyka, and R. Fellows, "Selection of topics for NICE technology appraisal 2005-2011: what matters most?," *Value in Health*, November 2014, *17* (7), A437–A438.

APPENDIX A. APPENDIX

A.1. **Product differentiation and price discrimination (Example).** To get intuition as to why an additional product or a new pack may be a source of increased revenue and how they may be substitutes, consider a simple example where a monopolist faces three consumers who differ in their valuation for a tablet and an extended release version (XR) due to differences in health conditions. Suppose further that additional packs of the tablet offer an instrument that allows for price discrimination (for simplicity, first degree in the example below). Assume also that marginal cost for either formulation is zero.

Consumer	Tablet (valuations)	Tablet XR (valuations)
A	150	100
B	100	150
C	10	100

In the baseline case, the monopolist has only one formulation, tablet, which she sells at a uniform price (i.e. no additional packs available). Then she would set a price of 100 and earn a revenue of TR = 200. Consumer C would be priced out of the market.

We now compare this benchmark to three cases: (i) a new XR formulation is available, (ii) additional packs for original tablet are available and hence price discrimination (PD) is possible, and (iii) both strategies can be employed (but additional packs are only for the original tablet).

(1) New XR. It is easy to see that by setting a price of 100 for the XR version, and increasing the price of the tablet to 150, the monopolist would earn TR = 350. There is both a market expansion effect (C now purchases the capsule) and a price increase effect (price of tablet increases by 50) and the TR changes by 350-200 = 150. With the additional product, the consumers with lower valuation for the original product (consumers B & C) switch to the new product where they have higher valuation, while consumer A stays with the original product.

(2) PD for original tab via additional packs: If the monopolist can perfectly price discriminate then the prices would be 150, 100 and 10 for the three consumers for a TR = 260. Relative to the base case, revenue increases by 60 and market expands (C is not priced out) though the average price decreases from 100 to (150+100+10)/3 = 86.67. Alternatively, if there was only limited scope for price discrimination, so that only two different prices can be offered (say there are only 2 pack variations available), then TR = 250, and relative to base case, the increase in revenue of 50 is via increase in average price (100+150)/2 = 125.

(3) New XR and PD for the original tablet. At a uniform price of 150 for XR (or an epsilon below that), consumer B would switch to it, while A and C will continue to buy original tablets at discriminatory prices of 150 and 10 for a TR= 310. Compared to the baseline case, TR changes by 310 − 200 = 110. However, this is not optimal pricing. If possible, the monopolist would prefer the low value consumer C to switch to XR as well, and can force them to do so by setting higher discriminatory price for C for tablet just above 10 and setting a lower uniform price of XR to 100. In this later case, her revenue would again increase to 350, which is same as that of the second scenario of only introducing an XR and demand does not expand.

This is obviously not an exhaustive example, but it is clear from it that introducing a new product or a pack can increase revenues, while doing both simultaneously may not necessarily be any better than just product differentiation and PD separately. With product differentiation the market is segmented, and low value consumers switch to the new product if they value it more. If there is

also PD in the original product, some low value consumers may not switch to the new product since they are already getting the product at their valuation (unless forced to do so for C as in case 4 above), in which case implementing both strategies simultaneously may not be optimal relatively to implementing them separately.

A.2. **Variable definitions and correlations.** Table A-1 provides a list of all variables used in the regressions.

TABLE A-1. Definitions

Variable	Definition
(1) $s_{b,t-1}$	Size: measured as ln sales (lagged one period).
(2) p_{bt}	Products: Count of total products (drugs) in the BU. A product is counted as a distinct product if either the molecule or the formulation (tablet, capsule, extended release etc.) changes. Alternative strengths or pack sizes are not counted as different products.
(3) v_{bt}	Pack Varieties: Total number of packs offered by the BU minus the number of products. Thus $v_{bt} = \#\text{packs}_{bt} - p_{bt}$ where #packs is count of items offered by the BU when either the molecule or formulation or strength or size differs. Example: If a BU has 4 different products and each comes in exactly one type of pack then $v_{bt} = 0$. If one of the 4 products is offered in two different strengths, then $v_{bt} = 1$.
(4) $p_{bt}v_{bt}$	Interaction of (2) and (3) above.
(5) $g_{b,t-1}$	Growth (lagged): Difference of log revenues over two consecutive periods $\ln R_{bt} - \ln R_{bt-1}$ where $\ln R_{b.}$ is log of revenue in two consecutive periods.
(6) x_{1bt}	PI: 1/0 dummy set to 1 if the BU has sales due to parallel imports.
(7) x_{2bt}	Generic: 1/0 dummy set to 1 if the BU has any sales due to generic drugs in its profile.
(8) x_{3bt}	Log price: Log of weighted prices of counting units in the BU. A counting unit is the smallest denomination of a drug, for instance an individual pill in a pack. The weights are derived from share of revenue associated with each pack in the BU.
(9) $x_{1c(t-1)}$	Products in class. Similar to p_{bt} but counted at class level (ATC4) rather than at the BU level.
(10) $x_{2c(t-1)}$	Pack varieties in class. Similar to v_{bt} but counted at class level (ATC4) rather than at the BU level.
(11) $x_{3c(t-1)}$	HHI at class level. Herfindahl index at class level computed from BU revenue of all BUs in the ATC4 class of the reference BU.
(12) $x_{4c(t-1)}$	Total number of firms in the ATC4 class of the reference BU.
(13) $\tau_{jt}, j = 1, \ldots, 13$	Set of dummy variables for each year and quarter.

The regressions also included lagged values of some of these variables.

TABLE A-2. Correlations (all variables in the main model)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
(1) $s_{b,t-1}$	1.000													
(2) p_{bt}	0.368	1.000												
(3) v_{bt}	0.425	0.732	1.000											
(4) $p_{bt}v_{bt}$	0.251	0.845	0.774	1.000										
(5) $g_{b,t-1}$	0.158	0.015	0.012	0.008	1.000									
(6) x_{1bt}	0.490	0.181	0.239	0.098	0.002	1.000								
(7) x_{2bt}	0.012	0.312	0.227	0.221	0.011	-0.173	1.000							
(8) x_{3bt}	0.108	-0.008	0.042	-0.017	0.028	0.007	-0.192	1.000						
(9) $x_{1c(t-1)}$	0.145	0.247	0.232	0.203	-0.016	0.097	-0.021	-0.173	1.000					
(10) $x_{2c(t-1)}$	0.173	0.213	0.332	0.198	-0.015	0.110	-0.012	-0.108	0.853	1.000				
(11) $x_{3c(t-1)}$	-0.229	-0.154	-0.157	-0.107	-0.011	-0.154	0.007	0.051	-0.475	-0.413	1.000			
(12) $x_{4c(t-1)}$	0.176	0.242	0.242	0.213	-0.014	0.114	0.001	-0.206	0.914	0.788	-0.590	1.000		
(13) $g_{b,t-2}$	0.121	0.009	0.007	0.005	-0.063	-0.004	0.012	0.041	-0.013	-0.013	-0.009	-0.012	1.000	
(14) $g_{b,t-3}$	0.064	0.004	0.004	0.004	-0.131	-0.010	-0.001	0.027	-0.011	-0.010	0.006	-0.013	-0.116	1.000
(15) $g_{b,t-4}$	0.043	0.001	0.001	0.001	-0.014	-0.017	-0.003	0.027	-0.008	-0.007	0.010	-0.011	-0.159	-0.083
(16) $p_{b,t-1}$	0.366	0.996	0.729	0.842	0.015	0.180	0.309	-0.009	0.248	0.212	-0.154	0.242	0.009	0.003
(17) $p_{b,t-2}$	0.363	0.993	0.726	0.839	0.008	0.179	0.305	-0.011	0.249	0.212	-0.153	0.243	0.008	0.003
(18) $p_{b,t-3}$	0.360	0.989	0.722	0.836	0.006	0.179	0.301	-0.013	0.250	0.211	-0.152	0.243	0.003	0.002
(19) $p_{b,t-4}$	0.357	0.985	0.719	0.833	0.006	0.178	0.298	-0.015	0.251	0.210	-0.151	0.244	0.001	-0.003
(20) $v_{b,t-1}$	0.424	0.732	0.996	0.773	0.009	0.240	0.226	0.040	0.233	0.331	-0.156	0.243	0.005	0.002
(21) $v_{b,t-2}$	0.421	0.732	0.992	0.773	0.003	0.240	0.224	0.038	0.234	0.329	-0.155	0.243	0.003	0.001
(22) $v_{b,t-3}$	0.418	0.731	0.988	0.772	0.000	0.240	0.223	0.036	0.234	0.327	-0.154	0.244	-0.003	-0.001
(23) $v_{b,t-4}$	0.415	0.731	0.983	0.772	0.000	0.240	0.221	0.033	0.234	0.325	-0.153	0.244	-0.004	-0.007
(24) $p_{b,t-1}v_{b,t-1}$	0.250	0.843	0.772	0.998	0.008	0.098	0.219	-0.018	0.203	0.197	-0.106	0.214	0.005	0.003
(25) $p_{b,t-2}v_{b,t-2}$	0.249	0.841	0.770	0.996	0.005	0.098	0.218	-0.018	0.203	0.196	-0.105	0.214	0.005	0.003
(26) $p_{b,t-3}v_{b,t-4}$	0.248	0.839	0.768	0.994	0.005	0.098	0.216	-0.019	0.203	0.195	-0.105	0.214	0.003	0.003
(27) $p_{b,t-4}v_{b,t-4}$	0.247	0.837	0.765	0.992	0.005	0.097	0.214	-0.019	0.203	0.193	-0.104	0.214	0.002	0.001

	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)	(26)	(27)
(15) $g_{b,t-4}$	1.000												
(16) $p_{b,t-1}$	0.000	1.000											
(17) $p_{b,t-2}$	-0.002	0.996	1.000										
(18) $p_{b,t-3}$	-0.001	0.993	0.996	1.000									
(19) $p_{b,t-4}$	-0.002	0.989	0.993	0.996	1.000								
(20) $v_{b,t-1}$	-0.001	0.731	0.728	0.725	0.721	1.000							
(21) $v_{b,t-2}$	-0.003	0.731	0.730	0.727	0.724	0.996	1.000						
(22) $v_{b,t-3}$	-0.004	0.731	0.730	0.729	0.727	0.993	0.996	1.000					
(23) $v_{b,t-4}$	-0.006	0.731	0.730	0.730	0.729	0.988	0.992	0.996	1.000				
(24) $p_{b,t-1}v_{b,t-1}$	0.001	0.843	0.841	0.838	0.835	0.774	0.774	0.774	0.774	1.000			
(25) $p_{b,t-2}v_{b,t-2}$	0.001	0.842	0.842	0.840	0.837	0.773	0.775	0.775	0.775	0.998	1.000		
(26) $p_{b,t-3}v_{b,t-4}$	0.000	0.840	0.841	0.841	0.839	0.771	0.773	0.776	0.776	0.996	0.998	1.000	
(27) $p_{b,t-4}v_{b,t-4}$	0.000	0.838	0.839	0.839	0.840	0.768	0.771	0.774	0.776	0.994	0.996	0.998	1.000

Correlations based on sample used in specification (4) in the main analysis. Year and quarter dummies omitted.

TABLE A-3. Correlations ($\Delta p_{b,(t-l)} \times \Delta v_{b,(t-l)}$)

		$\Delta p_{b,(t-l)}$				$\Delta v_{b,(t-l)}$			
		$t-1$	$t-2$	$t-3$	$t-4$	$t-1$	$t-2$	$t-3$	$t-4$
$\Delta p_{b,(t-l)}$	$t-1$	1.000							
	$t-2$	-0.034	1.000						
	$t-3$	0.024	-0.039	1.000					
	$t-4$	0.018	0.023	-0.045	1.000				
$\Delta v_{b,(t-l)}$	$t-1$	**0.276**	0.037	0.041	0.017	1.000			
	$t-2$	0.032	**0.268**	0.040	0.042	0.054	1.000		
	$t-3$	0.040	0.030	**0.263**	0.042	0.065	0.051	1.000	
	$t-4$	0.027	0.043	0.024	**0.266**	0.035	0.066	0.040	1.000

A.3. **Instruments.** To construct our instruments we make use of the fact that firms often operate in multiple ATC classes. In our data, a firm is on average operating in 9.48 different ATC4 classes, and hence it would appear as nine or ten different business units. Further, the raw correlation between the number of products produced by a given BU and the average number of products in all other BUs associated with the same firm is 0.251. Similarly, the correlation for pack varieties variable across same firm BUs is 0.226. We combine this observation with the assumption that the idiosyncratic error in the growth equation for a BU is serially uncorrelated and independent across ATC classes to construct our instruments for products, pack varieties, and interactions as follows. We compute the average value of number of products by the same firm in other related classes, where related classes are defined as business units within the same two-digit ATC classification. Specifically, we instrument Δp_{bt} with $\Delta \overline{p}_{-bt}$, where $\overline{p}_{-bt}$ is the firm's average value of p over time and over other BUs classes within the same ATC2 class, i.e. *excluding* the current BU and time period. Thus we obtain an instrument that varies by the business unit and time and which can be thought of as the deviation from the BU's long run steady state average number of products, where the latter is determined at the firm level in related classes. The raw correlation between p_{bt} and $\overline{p}_{-bt}$ is then 0.741. The logic is extended to derive instruments for the lagged values, i.e., $\Delta \overline{p}_{-b(t-1)}$ as instrument for $\Delta p_{b(t-1)}$, as well as for other lagged values of this variable. The instruments for number of pack varieties and its lagged values are constructed in a similar manner. We also use the total number of other business units the firm operates in, and its lagged values as additional instruments. Finally, the instrument for the interaction between number of products and varieties is constructed as the interaction of the instruments for number of products and those for number of varieties (and similarly for any lagged values of the interaction term).

Observe that if a firm does not operate in more than one class, we cannot construct an instrument for its products and varieties in the manner described above. However, to handle such cases, we can either drop the requirement of going outside the current ATC4 class and just use the average value from other time periods, or alternatively, use the average value from other time periods and any other ATC4 classes (i.e., do not restrict to average from the same ATC2 class assuming the firm operates in some other ATC class). Thus we also experimented with alternative ways of constructing these instruments by averaging the value of products or varieties over all other ATC4 classes, rather than just those within the same ATC2 class, or over other time periods except the current time period. The alternative instruments either turned out to be weak (when we went too wide and did not restrict to same ATC2 class), or were suspect for validity when we averaged only over other time periods. The main instruments listed here generally performed well when we restricted the sample to BUs that operate in more than one ATC4 class in various statistical tests relating to under and over identification, as well as first-stage weak instruments F-tests (see Online Appendix Table B-2 and Table B-3).

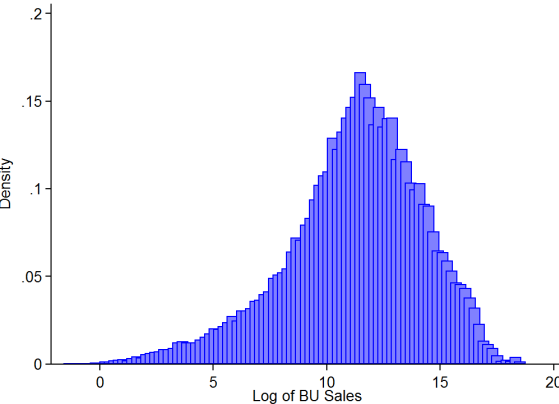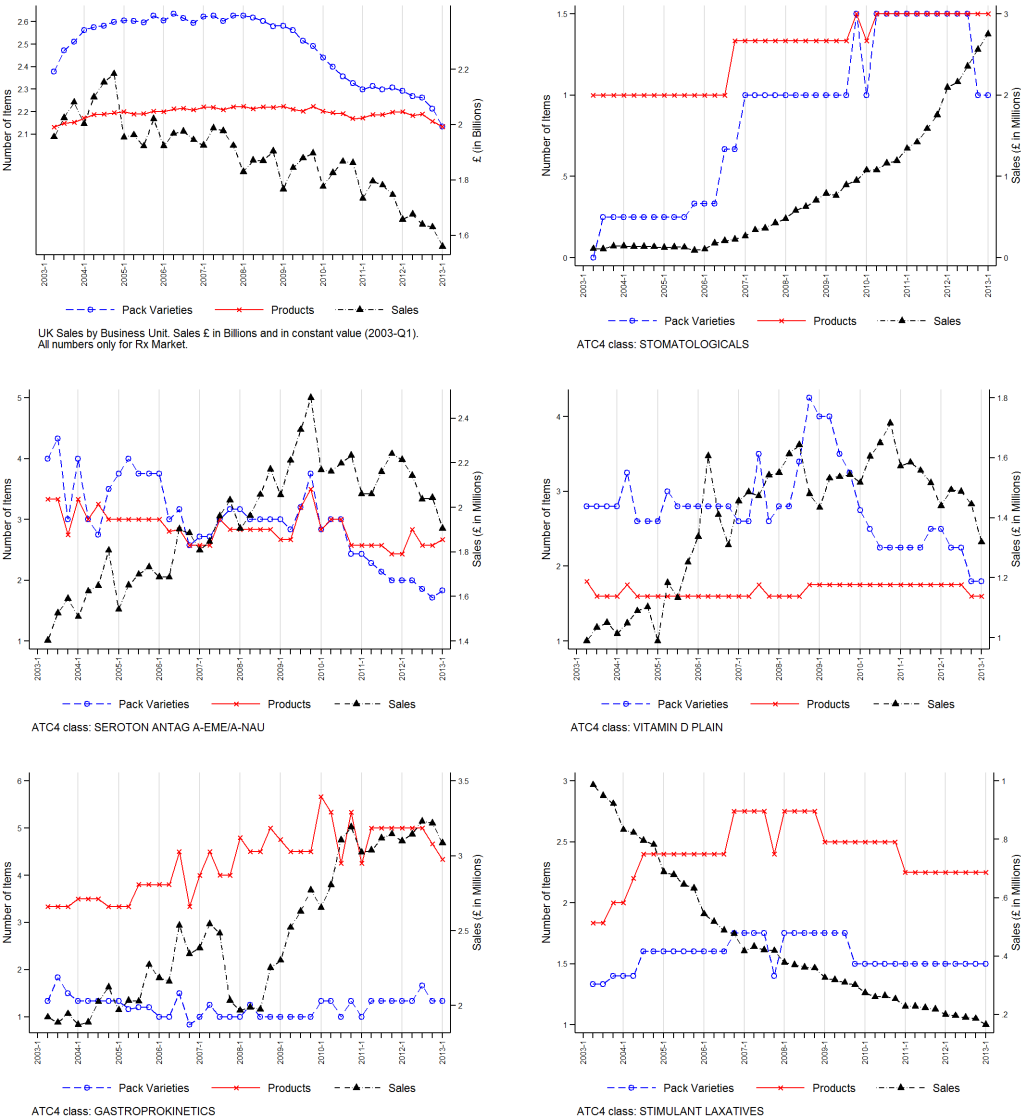FIGURE A-1. Distribution of Log Sales

FIGURE A-2. Sales, Products, and Pack Varieties – Overall and for selected ATC4 classes

A.4. **Specification (6) by Sub-samples.** This appendix provides estimates of the specification (6) by sub-samples. Only selected regression coefficients and marginal effects, along with first-stage statistics are shown.

TABLE A-4. First difference growth models (6) by sub-samples[†]

	(6)	(6S)	(6M)	(6L)	(6X)
Selected regression coefficients					
$s_{b,t-1}$: Lagged size	-0.71^a	-1.02^a	-0.65^a	-0.48^c	-0.86^a
	(0.12)	(0.20)	(0.22)	(0.25)	(0.18)
p_{bt}: Products	0.25^a	0.23	0.34^c	0.24^b	0.095^b
	(0.070)	(0.43)	(0.18)	(0.11)	(0.047)
v_{bt}: Varieties	0.18^a	0.28^c	0.29^a	0.11^a	0.068^a
	(0.035)	(0.16)	(0.083)	(0.042)	(0.015)
$p_{bt}v_{bt}$: Interaction	-0.80^a	0.044	-2.3^b	-0.57^b	-0.19^c
(Products×Varieties) ($\times 10^{-2}$)	(0.26)	(5.0)	(01.1)	(0.29)	(0.11)
$g_{b,t-1}$: Growth	-0.34^a	-0.17	-0.31	-0.37	-0.25
(Lag 1)	(0.11)	(0.17)	(0.22)	(0.23)	(0.16)
x_{3bt}: log Price	0.020	0.013	0.046	0.026	0.063^c
(Lag 2 of log wt. avg price)	(0.033)	(0.053)	(0.044)	(0.051)	(0.034)
Marginal effects of products and varieties [‡‡]					
Products (Short run)	0.105^a	0.125	0.194^c	0.130^b	0.047^b
	(0.029)	(0.199)	(0.105)	(0.059)	(0.024)
Products (Long run)	0.132^a	0.418^c	0.253^c	0.131^b	0.077^c
	(0.039)	(0.255)	(0.133)	(0.059)	(0.042)
Varieties (Short run)	0.071^a	0.152^a	0.158^a	0.058^a	0.034^a
	(0.013)	(0.048)	(0.042)	(0.020)	(0.007)
Varieties (Long run)	0.072^a	0.147^c	0.237^a	0.058^a	0.037^a
	(0.017)	(0.076)	(0.071)	(0.020)	(0.013)
Observations	23,035	5,612	8,326	9,097	20,999
R-squared	0.512	0.608	0.459	0.365	0.582
First Stage F-tests					
Endogenous Variable(s)					
$g_{b,t-1}$: Growth (Lag 1)	30.39	23.68	15.12	20.31	22.39
$p_{b,t}$: Products	69.33	13.00	32.35	191.28	58.81
$v_{b,t}$: Variety	218.1	98.10	86.60	137.30	214.1
$p_{bt} * v_{bt}$: Interaction	64.16	9.14	25.58	65.17	60.48
Under and Over Identification Tests					
Under-id $\chi^2(df)$	133.9 (2)	5.00 (2)	10.43 (2)	30.23 (2)	121.02 (2)
P-value	< .000	0.082	0.005	< .000	< 0.000
Over-id $\chi^2(df)$	1.54 (1)	0.03 (1)	0.59 (1)	1.80 (1)	0.398 (1)
P-vale	0.215	0.855	0.443	0.180	0.528

[†]Standard errors are in parenthesis and clustered by business unit (firm-ATC4 combination). Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively. All regressions include additional controls at the class and BU level, as well as indicator variables for year and quarter.

The results shown are for specification (6) in Table 4 which treats the first lag of growth as well as contemporaneous values of products, varieties and their interaction as endogenous. Specifications (6S), (6M), and (6L) are same as (6), but on sub-samples by initial size of business unit being (s)mall, (m)edium and (l)arge, while (6X) restricts (6) to BUs that do not exit the sample.

[‡‡] Marginal effects are computed at the sample mean. They account for the interaction terms, the lagged values of the variable as well as the lagged values of growth. Standard errors are computed using the delta method.

Appendix B. Online appendix

B.1. **Regression coefficients.** A shorter version of this table is included in the main paper (see Table 4) where it shows the coefficient of selected variables. This appendix contains regression coefficients on all the variables.

Table B-1. FD Growth Models[†]

	OLS			IV			(4) By sub-samples (IV)			
	(1)	(2)	(3)	(4)	(5)	(6)	(4S)	(4M)	(4L)	(4X)
$s_{b,t-1}$: Size (Lagged Size of Bus. Unit)	-1.06^a (0.016)	-1.05^a (0.016)	-1.05^a (0.034)	-0.76^a (0.099)	-0.71^a (0.12)	-0.71^a (0.12)	-0.77^a (0.13)	-0.77^a (0.18)	-1.12^a (0.35)	-0.79^a (0.10)
p_{bt}: Products	0.28^a (0.041)	0.30^a (0.044)	0.27^a (0.043)	0.26^a (0.044)	0.25^a (0.070)	0.25^a (0.070)	0.58^a (0.10)	0.40^a (0.12)	0.15^b (0.057)	0.15^a (0.024)
v_{bt}: Variety	0.16^a (0.018)	0.15^a (0.018)	0.14^a (0.019)	0.14^a (0.019)	0.17^a (0.034)	0.18^a (0.035)	0.32^a (0.053)	0.25^a (0.055)	0.079^a (0.021)	0.074^a (0.012)
$p_{bt}v_{bt}$: Interaction (Products×Variety) ($\times 10^{-2}$)	-0.70^a (0.13)	-0.70^a (0.13)	-0.65^a (0.12)	-0.63^a (0.12)	-0.77^a (0.26)	-0.80^a (0.26)	-3.40^a (0.90)	-1.60^a (0.54)	-0.30^a (0.11)	-0.31^a (0.064)
$g_{b,t-1}$: Growth (Lag 1)			-0.0030 (0.026)	-0.28^a (0.098)	-0.34^a (0.11)	-0.34^a (0.11)	-0.36^a (0.12)	-0.19 (0.17)	0.22 (0.34)	-0.29^a (0.098)
x_{1bt}: Any PI Sales? (1/0 Dummy, 1 if true)	0.12^a (0.026)	0.12^a (0.026)	0.12^a (0.025)	0.11^a (0.026)	0.13^a (0.035)	0.13^a (0.035)	0.16^c (0.085)	0.11^a (0.037)	0.081^b (0.032)	0.099^a (0.023)
x_{2bt}: Any generic products? (1/0 Dummy, 1 if true)	-0.0039 (0.11)	-0.099 (0.088)	-0.071 (0.095)	-0.057 (0.098)	0.031 (0.12)	0.034 (0.12)	-0.41 (0.44)	0.24 (0.22)	-0.13^b (0.052)	-0.049 (0.071)
x_{3bt}: Log Price (Lag 2 Wt avg price of BU drugs)	0.0072 (0.027)	-0.0016 (0.024)	0.050^b (0.024)	0.027 (0.024)	0.017 (0.034)	0.020 (0.033)	-0.026 (0.028)	0.12^c (0.068)	0.052 (0.032)	0.040^c (0.023)
$x_{1c,t-1}$: Products in Class ($\times 10^{-2}$) (Lag 1 of Products by Others)	-0.53 (0.42)	-0.63 (0.45)	-0.71 (0.46)	-0.59 (0.46)	-0.37 (0.71)	-0.36 (0.71)	-0.54 (1.30)	0.085 (0.940)	-1.20^b (0.510)	-0.74^c (0.390)
$x_{2c,t-1}$: Varieties in Class ($\times 10^{-2}$) (Lag 1 of Varieties by Others)	-0.11 (0.19)	-0.033 (0.19)	0.014 (0.20)	-0.019 (0.19)	-0.46 (0.34)	-0.42 (0.34)	0.41 (0.48)	0.48 (0.45)	-0.32^c (0.19)	0.09 (0.18)
$x_{3c,t-1}$: HHI (Lag 1)	0.34^b (0.17)	0.30^c (0.18)	0.41^b (0.18)	0.41^b (0.18)	0.30 (0.22)	0.30 (0.22)	0.28 (0.18)	0.68 (0.45)	0.25 (0.46)	0.40^b (0.17)
$x_{4c,t-1}$: Firms in Class (Lag 1) ($\times 10^{-2}$)	0.87 (0.92)	0.61 (0.90)	0.57 (0.91)	0.30 (0.92)	-0.10 (1.20)	-0.061 (1.30)	0.96 (2.50)	-1.30 (1.70)	0.80 (1.10)	0.047 (0.790)
(Lags)										
Observations	48,799	44,856	42,994	42,994	23,035	23,035	12,567	14,815	15,612	38,345
R-squared	0.510	0.507	0.508	0.502	0.512	0.512	0.556	0.461	0.425	0.560

[†]Standard errors are in parenthesis and clustered by business unit (firm-ATC4 combination). Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively.

Specification (1) does not include any lags, (2) is a distributed lag model and includes up to four lags of products, varieties and interactions. Specification (3) is an augmented distributed lag model that also includes four lags of the dependent variable (i.e., growth). Specification (4) treats only the first lag of growth as endogenous, (5) treats the first lag of growth, as well as products, and varieties, and their interaction and all four lags of each of these as endogenous variables (total 16), and specification (6) treats the first lag of growth and only the contemporaneous values of products, varieties and interaction as endogenous variables (total of 4 endogenous variables). Specifications (4S), (4M), and (4L) are same as (4), but on sub-samples by initial size of business unit being (s)mall, (m)edium and (l)arge, while (4X) restricts (4) to BUs that do not exit the sample.

TABLE B-1. FD Growth Models[†]

	OLS			IV			(4) By sub-samples (IV)			
	(1)	(2)	(3)	(4)	(5)	(6)	(4S)	(4M)	(4L)	(4X)
$g_{b,t-2}$: Growth (Lag 2)			-0.13^a (0.033)	-0.34^a (0.087)	-0.42^a (0.093)	-0.42^a (0.092)	-0.41^a (0.12)	-0.26^b (0.13)	0.059 (0.24)	-0.37^a (0.086)
$g_{b,t-3}$: Growth (Lag 3)			-0.15^a (0.033)	-0.28^a (0.065)	-0.35^a (0.072)	-0.35^a (0.072)	-0.33^a (0.093)	-0.21^b (0.088)	-0.029 (0.15)	-0.32^a (0.066)
$g_{b,t-4}$: Growth (Lag 4)			-0.019 (0.012)	-0.076^a (0.025)	-0.11^a (0.030)	-0.11^a (0.030)	-0.11^a (0.031)	-0.018 (0.052)	0.063 (0.065)	-0.082^a (0.027)
$p_{b,t-1}$: Products (Lag1)		0.10^a (0.021)	0.089^a (0.021)	0.081^a (0.021)	0.042 (0.032)	0.077^a (0.026)	0.22^b (0.091)	0.15^b (0.070)	0.019 (0.017)	0.074^a (0.021)
$p_{b,t-2}$: Products (Lag 2)		0.030 (0.025)	0.059^b (0.023)	0.044^b (0.022)	0.11^b (0.056)	0.070^b (0.029)	0.14^b (0.061)	0.058 (0.060)	0.028 (0.030)	0.031^c (0.017)
$p_{b,t-3}$: Products (Lag 3)		-0.018 (0.016)	-0.00073 (0.016)	-0.022 (0.019)	-0.033 (0.025)	-0.037^c (0.021)	-0.033 (0.069)	-0.015 (0.053)	0.0095 (0.019)	-0.026 (0.016)
$p_{b,t-4}$: Products (Lag 4)		0.050 (0.031)	-0.00099 (0.019)	-0.028 (0.023)	-0.021 (0.031)	-0.044^c (0.024)	-0.058 (0.082)	0.038 (0.053)	0.0038 (0.018)	-0.041^c (0.022)
$v_{b,t-1}$: Variety (Lag 1)		0.033^a (0.0088)	0.029^a (0.0086)	0.021^b (0.0088)	0.020 (0.014)	0.025^b (0.012)	0.048 (0.033)	0.082^a (0.031)	0.010 (0.0090)	0.022^a (0.0075)
$v_{b,t-2}$: Variety (Lag 2)($\times 10^{-2}$)		0.90 (0.79)	1.80^b (0.89)	0.66 (0.90)	1.90 (2.30)	0.082 (1.20)	2.80 (2.80)	2.40 (2.20)	1.50 (1.30)	0.50 (0.61)
$v_{b,t-3}$: Variety (Lag 3) ($\times 10^{-2}$)		0.55 (0.61)	0.42 (0.65)	-0.79 (0.81)	-1.40 (1.10)	-1.40^c (0.85)	0.40 (3.40)	0.30 (1.90)	0.17 (0.79)	-1.20 (0.73)
$v_{b,t-4}$: Variety (Lag 4) ($\times 10^{-2}$)		1.90^a (0.66)	0.025 (0.66)	-1.20 (0.80)	0.14 (1.20)	-0.68 (0.88)	-2.10 (3.20)	0.48 (2.00)	0.66 (0.73)	-1.70^b (0.70)
$p_{b,t-1}v_{3,t-1}$: Interaction (Products×Variety, Lag 1) ($\times 10^{-2}$)		-0.23^a (0.054)	-0.20^a (0.051)	-0.17^a (0.049)	-0.060 (0.12)	-0.21^a (0.075)	-1.20 (0.77)	-0.88^c (0.51)	-0.034 (0.037)	-0.14^a (0.043)
$p_{2,t-2}v_{3,t-2}$: Interaction (Products×Variety, Lag 2) ($\times 10^{-2}$)		-0.082^c (0.047)	-0.13^a (0.049)	-0.090^b (0.045)	-0.24 (0.17)	-0.13^c (0.064)	-0.60 (0.59)	-0.56^a (0.20)	-0.062 (0.060)	-0.070^b (0.033)
$p_{2,t-3}v_{3,t-3}$: Interaction (Products×Variety, Lag 3) ($\times 10^{-2}$)		0.036 (0.037)	0.0071 (0.037)	0.062 (0.043)	0.082 (0.084)	0.10^c (0.054)	-0.57 (0.63)	0.051 (0.27)	-0.0046 (0.038)	0.068^c (0.036)
$p_{2,t-4}v_{3,t-4}$: Interaction (Products×Variety, Lag 4) ($\times 10^{-2}$)		-0.11^c (0.065)	0.0035 (0.042)	0.064 (0.049)	-0.059 (0.093)	0.078 (0.057)	-0.51 (0.67)	-0.23 (0.22)	-9.0e-04 (0.033)	0.087^b (0.042)
τ_{1t}: Yr 2004	0.0056			-0.014	-0.068^b	-0.066^b	0.020	-0.048	0.0083	0.025
Observations	48,799	44,856	42,994	42,994	23,035	23,035	12,567	14,815	15,612	38,345
R-squared	0.510	0.507	0.508	0.502	0.512	0.512	0.556	0.461	0.425	0.560

[†]Standard errors are in parenthesis and clustered by business unit (firm-ATC4 combination). Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively.

Specification (1) does not include any lags, (2) is a distributed lag model and includes up to four lags of products, varieties and interactions. Specification (3) is an augmented distributed lag model that also includes four lags of the dependent variable (i.e., growth). Specification (4) treats only the first lag of growth as endogenous, (5) treats the first lag of growth, as well as products, and varieties, and their interaction and all four lags of each of these as endogenous variables (total 16), and specification (6) treats the first lag of growth and only the contemporaneous values of products, varieties and interaction as endogenous variables (total of 4 endogenous variables). Specifications (4S), (4M), and (4L) are same as (4), but on sub-samples by initial size of business unit being (s)mall, (m)edium and (l)arge, while (4X) restricts (4) to BUs that do not exit the sample.

Table B-1. FD Growth Models[†]

	OLS			IV			(4) By sub-samples (IV)			
	(1)	(2)	(3)	(4)	(5)	(6)	(4S)	(4M)	(4L)	(4X)
(1/0 Dummy, 1 if true)	(0.012)			(0.023)	(0.031)	(0.032)	(0.063)	(0.037)	(0.029)	(0.023)
τ_{2t}: Yr 2005 (1/0 Dummy, 1 if true)		-0.029^c (0.016)	-0.016 (0.018)	-0.028 (0.018)	-0.071^a (0.022)	-0.070^a (0.022)	-0.0042 (0.036)	-0.047 (0.032)	-0.024 (0.027)	0.0086 (0.017)
τ_{3t}: Yr 2006 (1/0 Dummy, 1 if true)	0.017 (0.013)	-0.019 (0.016)	-0.0063 (0.017)	-0.021 (0.019)	-0.071^a (0.023)	-0.070^a (0.023)	-0.018 (0.040)	-0.044 (0.033)	-0.0093 (0.024)	0.038^b (0.018)
τ_{4t}: Yr 2007 (1/0 Dummy, 1 if true)	0.023^b (0.011)	-0.011 (0.014)	0.00097 (0.017)	-0.017 (0.018)	-0.074^a (0.023)	-0.072^a (0.023)	-0.011 (0.036)	-0.027 (0.032)	-0.012 (0.026)	0.016 (0.017)
τ_{5t}: Yr 2008 (1/0 Dummy, 1 if true)	0.023^c (0.012)	-0.011 (0.015)	-0.0018 (0.017)	-0.016 (0.018)	-0.069^a (0.023)	-0.067^a (0.023)	0.017 (0.039)	-0.037 (0.031)	-0.023 (0.025)	0.021 (0.017)
τ_{6t}: Yr 2009 (1/0 Dummy, 1 if true)	0.0037 (0.012)	-0.026^c (0.015)	-0.010 (0.017)	-0.025 (0.018)	-0.077^a (0.022)	-0.075^a (0.022)	-0.0099 (0.037)	-0.058^c (0.032)	-0.013 (0.023)	0.0099 (0.017)
τ_{7t}: Yr 2010 (1/0 Dummy, 1 if true)	0.0083 (0.011)	-0.019 (0.015)	-0.0050 (0.017)	-0.017 (0.018)	-0.071^a (0.024)	-0.070^a (0.024)	-0.0031 (0.035)	-0.033 (0.033)	-0.0099 (0.026)	0.012 (0.017)
τ_{8t}: Yr 2011 (1/0 Dummy, 1 if true)	-0.0097 (0.011)	-0.035^b (0.015)	-0.016 (0.016)	-0.030^c (0.018)	-0.075^a (0.023)	-0.074^a (0.023)	-0.012 (0.037)	-0.040 (0.032)	-0.030 (0.026)	0.0046 (0.017)
τ_{9t}: Yr 2012 (1/0 Dummy, 1 if true)	0.022^c (0.012)	-0.0098 (0.014)	0.0021 (0.017)	-0.0086 (0.018)	-0.058^b (0.024)	-0.055^b (0.024)	0.030 (0.036)	-0.039 (0.032)	-0.017 (0.025)	0.020 (0.017)
τ_{10t}: Yr 2013 (1/0 Dummy, 1 if true)	0.022 (0.018)	-0.0015 (0.022)	0.012 (0.023)							
τ_{11t}: Qtr 2 (1/0 Dummy, 1 if true)	0.065^a (0.0087)	0.081^a (0.0088)	0.072^a (0.0084)	0.071^a (0.0084)	0.081^a (0.012)	0.081^a (0.012)	0.031^c (0.017)	0.085^a (0.016)	0.088^a (0.012)	0.060^a (0.0071)
τ_{12t}: Qtr 3 (1/0 Dummy, 1 if true)	0.065^a (0.013)	0.055^a (0.012)	0.042^a (0.0081)	0.043^a (0.0082)	0.045^a (0.011)	0.045^a (0.011)	0.045^b (0.018)	0.044^a (0.016)	0.027^b (0.013)	0.037^a (0.0069)
τ_{13t}: Qtr 4 (1/0 Dummy, 1 if true)	0.053^a (0.011)	0.053^a (0.011)	0.050^a (0.0087)	0.050^a (0.0088)	0.060^a (0.012)	0.060^a (0.012)	0.039^c (0.021)	0.071^a (0.015)	0.033^a (0.011)	0.040^a (0.0075)
x_{0t}: Constant	-0.094^a (0.011)	-0.070^a (0.014)	-0.082^a (0.017)	-0.067^a (0.015)	-0.019 (0.018)	-0.020 (0.018)	-0.084^a (0.030)	-0.064^b (0.027)	-0.058^b (0.025)	-0.072^a (0.015)
Observations	48,799	44,856	42,994	42,994	23,035	23,035	12,567	14,815	15,612	38,345
R-squared	0.510	0.507	0.508	0.502	0.512	0.512	0.556	0.461	0.425	0.560

[†]Standard errors are in parenthesis and clustered by business unit (firm-ATC4 combination). Superscripts a, b, c indicate significance at 1%, 5% and 10%, respectively.

Specification (1) does not include any lags, (2) is a distributed lag model and includes up to four lags of products, varieties and interactions. Specification (3) is an augmented distributed lag model that also includes four lags of the dependent variable (i.e., growth). Specification (4) treats only the first lag of growth as endogenous, (5) treats the first lag of growth, as well as products, and varieties, and their interaction and all four lags of each of these as endogenous variables (total 16), and specification (6) treats the first lag of growth and only the contemporaneous values of products, varieties and interaction as endogenous variables (total of 4 endogenous variables). Specifications (4S), (4M), and (4L) are same as (4), but on sub-samples by initial size of business unit being (s)mall, (m)edium and (l)arge, while (4X) restricts (4) to BUs that do not exit the sample.

Table B-2. First stage and other tests for instruments

	(4)	(5)	(6)	(4S)	(4M)	(4L)	(4X)
			Weak instruments tests				
Endogenous Variable(s)			First-stage F-statistics				
$g_{b,t-1}$: Growth (Lag 1)	59.42	13.45	30.39	37.97	15.19	15.19	31.18
$p_{b,t}$: Products		37.80	69.33				
$v_{b,t}$: Variety		80.94	218.1				
$p_{bt} * v_{bt}$: Interaction		32.26	64.16				
$p_{b,t-1}$: Products (Lag 1)		64.98					
$p_{b,t-2}$: Products (Lag 2)		86.81					
$p_{b,t-3}$: Products (Lag 3)		84.42					
$p_{b,t-4}$: Products (Lag 4)		91.59					
$v_{b,t-1}$:Variety (Lag1)		76.66					
$v_{b,t-2}$:Variety (Lag2)		74.63					
$v_{b,t-3}$:Variety (Lag3)		70.51					
$v_{b,t-4}$:Variety (Lag4)		72.31					
$p_{b,t-1} * v_{b,t-1}$: Interaction (Lag 1)		37.18					
$p_{b,t-2} * v_{b,t-2}$: Interaction (Lag 2)		55.30					
$p_{b,t-3} * v_{b,t-3}$: Interaction (Lag 3)		39.25					
$p_{b,t-4} * v_{b,t-4}$: Interaction (Lag 4)		34.47					
			Under and Over Identification Tests				
(Ho: Model is under identified)							
Under-id $\chi^2(df)$	169.01 (1)	139.6 (6)	133.9 (2)	92.62 (1)	56.59 (1)	22.28 (1)	121.54 (1)
P-value	< .000	< .000	< .000	< .000	< .000	< .000	< .000
(Ho: Model is over identified)							
Over-id $\chi^2(df)$	.	6.06 (5)	1.54 (1)	.	.	.	.
P-value	.	0.306	0.215	.	.	.	.

TABLE B-3. Instruments test w./ alternative clustering

Clustering at the Firm level

Endogenous Variable(s)	(4)	(5)	(6)	(4S)	(6M)	(6L)	(6X)
			First-stage F-statistics				
$g_{b,t-1}$: Growth (Lag 1)	34.06	59.49	36.17	34.13	12.08	14.58	14.84
$p_{b,t}$: Products		926.4	211.4				
$v_{b,t}$: Variety		3804	131.1				
$p_{bt} * v_{bt}$: Interaction		1426	172.1				
$p_{b,t-1}$: Products (Lag 1)		1701					
$p_{b,t-2}$: Products (Lag 2)		1663					
$p_{b,t-3}$: Products (Lag 3)		2762					
$p_{b,t-4}$: Products (Lag 4)		1094					
$v_{b,t-1}$:Variety (Lag1)		9204					
$v_{b,t-2}$:Variety (Lag2)		5728					
$v_{b,t-3}$:Variety (Lag3)		808.9					
$v_{b,t-4}$:Variety (Lag4)		505.6					
$p_{b,t-1} * v_{b,t-1}$: Interaction (Lag 1)		1515					
$p_{b,t-2} * v_{b,t-2}$: Interaction (Lag 2)		1828					
$p_{b,t-3} * v_{b,t-3}$: Interaction (Lag 3)		5879					
$p_{b,t-4} * v_{b,t-4}$: Interaction (Lag 4)		3677					

Under and Over Identification Tests

Ho: Model is under identified							
Under-id $\chi^2(df)$	44.10 (1)	25.27 (6)	17.18 (2)	33.66 (1)	27.08 (1)	13.58 (1)	36.32 (1)
P-value	< .000	< .000	< .000	< .000	< .000	< .000	< .000
Ho: Model is over identified							
Over-id $\chi^2(df)$	.	7.23 (5)	0.48 (1)	.	.	.	.
P-value	.	0.204	0.490	.	.	.	.

Clustering at the ATC4 level

Endogenous Variable(s)	(4)	(5)	(6)	(4S)	(4M)	(4L)	(4X)
			First-stage F-statistics				
$g_{b,t-1}$: Growth (Lag 1)	54.33	21.40	23.89	30.51	16.17	15.17	29.95
$p_{b,t}$: Products		49.31	48.58				
$v_{b,t}$: Variety		88.42	198.89				
$p_{bt} * v_{bt}$: Interaction		90.75	54.19				
$p_{b,t-1}$: Products (Lag 1)		80.49					
$p_{b,t-2}$: Products (Lag 2)		90.57					
$p_{b,t-3}$: Products (Lag 3)		71.44					
$p_{b,t-4}$: Products (Lag 4)		71.57					
$v_{b,t-1}$:Variety (Lag1)		109.8					
$v_{b,t-2}$:Variety (Lag2)		79.19					
$v_{b,t-3}$:Variety (Lag3)		71.93					
$v_{b,t-4}$:Variety (Lag4)		77.17					
$p_{b,t-1} * v_{b,t-1}$: Interaction (Lag 1)		104.6					
$p_{b,t-2} * v_{b,t-2}$: Interaction (Lag 2)		127.4					
$p_{b,t-3} * v_{b,t-3}$: Interaction (Lag 3)		73.21					
$p_{b,t-4} * v_{b,t-4}$: Interaction (Lag 4)		50.32					

Under and Over Identification Tests

(Ho: Model is under identified)							
Under-id $\chi^2(df)$	77.82 (1)	79.79 (6)	77.88(2)	61.62 (1)	42.99 (1)	18.71 (1)	62.30 (1)
P-value	< .000	< .000	< .000	< .000	< .000	< .000	< .000
(Ho: Model is over identified)							
Over-id $\chi^2(df)$	.	5.92 (5)	1.58 (1)	.	.	.	.
P-value	.	0.314	0.209	.	.	.	.