Demand Forces of Technical Change Evidence from the Chinese Manufacturing Industry*†

Andreas Beerli, Franziska J. Weiss, Fabrizio Zilibotti and Josef Zweimüller‡

November 30, 2012

Abstract

This paper investigates the effect of market size on innovation activities across different durable good industries in the Chinese manufacturing sector. We use a potential market size measure driven only by changes in the Chinese income distribution which is exogenous to changes in prices and qualities of durable goods to instrument for actual future market size. Results indicate that an increase in market size by one percentage point leads to an increase of 4.4% in R&D inputs, an increase in labour productivity by 6.5% and an increase in the likelihood of a successful product innovation by about 1.1 percentage points. These findings are robust controlling for export behaviour of firms and supply side drivers of R&D.

Keywords: China, Demand-induced Innovation, Economic Growth, Market Size, Consumer Preferences.

JEL classification: D21, L16, L60, O31, O33.

*!!! PRELIMINARY, PLEASE DO NOT CITE WITHOUT PERMISSION !!!

†We thank Berthold Herrendorf, Christian Helmers, Guy Michaels, Uwe Sunde, Nico Voigtländer and Xiaodong Zhu for their fruitful suggestions. Further we thank all participants at the Zurich Workshop in Economics 2012, participants at the DEGIT conference 2012 and seminar participants at the University of Zurich for their helpful comments.

‡University of Zurich, Department of Economics, Mühlebachstrasse 86, CH-8008 Zürich. E-mails: andreas.beerli@econ.uzh.ch; franziska.weiss@econ.uzh.ch, fabrizio.zilibotti@econ.uzh.ch, josef.zweimueller@econ.uzh.ch.

1 Introduction

What is the effect of the rapidly growing middle class in China on innovation activities of Chinese manufacturing firms? In this paper we would like to analyse the interplay between the rapidly growing middle class of new consumers and the process of technical change in the Chinese economy. In China, unprecedented in modern times, average income grew by roughly 10 percent p.a. since the outset of the reforms in 1978, lifting over half a billion people out of poverty and creating a soaring middle class which accounts for over half of the population in recent population surveys.[1]

Dating back to Engel (1857), it is one of the most robust empirical findings in economics that consumers change the composition of their consumption bundle as they become richer. Growth in their incomes induce them to reallocate relatively less expenditure to necessity goods and more to luxury goods, thus creating markets for new consumer goods. The idea that market size and profit incentives are the key determinant of innovation activities of firms was first articulated by Schmookler (1966) in his early study on inventions and growth:

"It is that (1) invention is largely an economic activity which, like other economic activities, is pursued for gain;" (Schmookler, 1966: 206).

Since then, this argument became widely established in the endogenous growth literature, where the rate of technological progress is largely determined by profit opportunities.[2] In particular, directed technical change models emphasize the positive influence of market size on the *direction* and *bias* of research.[3]

However, until recently the actual structure of the demand, resulting from heterogeneous consumers with preferences in the spirit of Engel, played a relatively minor role in models of structural change and growth. Kongsmut, Rebelo and Xie (2001) were among the first to develop a theoretical framework explaining structural change driven by an income effect. Yet, their model allows for a balanced growth path only under some knife-edge condition on the preference structure. Boppart (2011) relies on a more general form of non-Gorman preferences generating structural change within the economy while allowing for a balanced growth path on the aggregate. With a non-unitary expenditure elasticity of demand, the non-homotheticity of preferences implies that an increase in real per-capita income levels affect the sectoral expenditure shares which, in turn, change the structure of durable

[1]Cf. Figure 3 in the next sections. Income groups of CHNS data, classification according to Worldbank (2009).

[2]Compare, for example, Aghion and Howitt (1992) or Grossman and Helpman (1991).

[3]The idea of induced innovations was already formalized in Habakkuk (1962) and Schmookler (1966) (among others). Compare Acemoglu (1998) and Acemoglu (2002) for the idea of directed technical change.

good ownership.

In a related strand of the literature on demand-induced innovation, Foellmi and Zweimüller (2006) build a theoretical model to analyse the link between heterogeneous consumers (with respect to their initial wealth) with non-homothetic preferences and the pricing and innovation decision of monopolistic firms. As consumers have hierarchic preferences, newly invented products are initially demanded only by the rich households, while the poor only consume necessities. As incomes grow, formerly poor consumers start demanding more luxuries which triggers a shift in good specific market size and product innovation incentives. Foellmi, Wuergler and Zweimüller (2009) extend this framework to differentiate explicitly between product innovation and process innovation. They find that a substantial reduction in the number of poor people (accompanied by a major drop in inequality in their paper) may induce a period of industrial change where innovation is directed towards process innovation.

Although theoretically appealing, empirical evidence along these lines is relatively scarce. Two notable exceptions are Acemoglu and Linn (2004) and Boppart and Weiss (2012). Acemoglu and Linn (2004) document a causal link between market size and innovation rates within the U.S. pharmaceutical sector. Using the ongoing demographic change as exogenous source of variation in market size for different drug categories, they find positive effects of market size on innovation across different drug categories. While their study targets the pharmaceutical industry, Boppart and Weiss (2012) extend their set-up to a comprehensive cross-industry analysis. Using industrial TFP and R&D input data for the U.S. between 1977-2007, they find a significant positive effect of a sector's market share, on sector specific R&D investments and productivity growth rates.

In this paper, we reconcile the views of these different strands of the literature linking the Chinese structural transformation and innovation activities across different manufacturing sectors to changes in market size driven by shifts in the Chinese income distribution. In particular, we exploit variation in durable good ownership rates of Chinese households to identify innovation activities across different manufacturing sectors. For this reason, we link the flow measure of new durable good acquisitions of households surveyed in the China Health and Nutrition Survey (CHNS) to manufacturing firms from the Annual Survey of Industrial Production (ASIP). As the observed, actual market size of durable good industries is likely to be endogenous, with qualitatively improved or cheaper products having larger markets, we use an instrumental variables strategy to account for this. Particularly, we follow Acemoglu and Linn (2004) and construct a measure of potential market size using the income and ownership information from the CHNS to construct instruments for the observed actual

market size. In so doing, we fix income group specific ownership to a particular base-year and use the changing population shares of income groups across time to calculate a measure of potential market size in other years. Then, the time dynamic of this measure of potential market size is only driven by changes in the income distribution and not by changes in ownership patterns of a given income group which might be induced by changes prices or quality of goods.
The the rich firm-level data set allows to evaluate the effect of changes in market size on different measures of firm-level innovation. Particularly, we look at the likelihood of positive new product sales, a measure for product innovation, at firm-level investment flows as a proxy for R&D investments and at labour productivity as a standard performance measure of firms' being influenced by both product and process innovation.[4]

The results of our instrumental variables estimation suggest that a one percentage point increase in market size over the next five years raises the probability of a successful product innovation by about 1.1 percentage points, increases labour productivity by 6.5% and leads to a 4.4% increase in R&D inputs. These results were found to be robust to including a rich set of firm-level determinants of R&D and the sector market concentration. As China's economy is still heavily export driven, a major concern is that the non-domestic demand might be an important driver of technical change. We test the robustness of our results controlling for firm specific export behaviour on the intensive and extensive margin. Unsurprisingly, we can show that the domestic market size effect is weaker for firms heavily involved in trade. Furthermore, we can show that our results are robust to supply side drivers of R&D affecting innovation opportunities of Chinese firms by including a measure of worldwide technology potential reported by Swiss firms in our baseline regression.

Our findings relate to a growing literature on the role of domestic demand in the process of technical progress in developing China. Brandt, Rawski and Sutton (2008), for instance, describe the rapid expansion of new consumer good industries in China as a response to a surge in domestic demand arising from increased household incomes. Along these lines, Hu and Jefferson (2008) argue that the increase in China's R&D intensity and patenting activity in the recent decade might be influenced to a major extend by an increasing demand for technology intensive goods.

The rest of the paper is organized as follows. Section 2 describes our data sources, provides some descriptive statistics and explains our empirical strategy. Section 3 presents all empirical results and different robustness checks. Finally, Section 4 concludes.

[4]Labour productivity was shown to be far less noisy than TFP and thus less prone to measurement error.

2 Data, Descriptive Statistics and Empirical Strategy

In this paper, we want to establish a causal relationship between the changes in aggregate demand patterns for different kinds of durable goods driven by growth in household income and shifting innovation activities across different Chinese manufacturing sectors. For this purpose we link micro-data of manufacturing firms to an aggregate market size measure taken from household survey data.

2.1 Actual Market Size

To construct a market size measure which is driven only by quasi-exogenous shifts in the income distribution we exploit micro-level information on Chinese households' durable good ownership[5] from the China Health and Nutrition Survey (CHNS). The CHNS was collected collected in eight waves between 1989 and 2009. The survey covers a representative sample of Chinese urban and rural households across nine provinces with substantial variation in geography, economic development and public resources. In contrast to the NBS households survey, the CHNS micro data are publicly available and are widely used in the literature.[6]

We measure actual market size or sales of different durable goods in the following way. From the CHNS we observe the number of items owned per household of durable good j at time t. The stock of durable good j at time t per household, $Stock_{j,t}^{actual}$, is simply the average number of items owned per household. In this framework, an new item of a particular durable good is acquired by a household (and hence sold by a firm) through one of the following three channels: First, if a household becomes a new owner and acquires a particular durable good for the first time. Second, when an existing owner buys an additional item of the same durable good. And third, if a household replaces a worn out item (replacement demand). In the CHNS, the first two channels[7] of sales are captured by

[5]Working with durable goods ownership rather than household expenditure data has some important advantages but also bears some difficulties. The main advantage is that CHNS' coverage of a relatively broad set of different durable goods allows to construct a market size measure with substantial sector and time variation which can be linked relatively straightforward to different sectors in the manufacturing data. Second, the lumpy nature of durable goods creates an interesting variation in ownership profiles across the income distribution which can be exploited to create an exogenous measure of market size. As a major disadvantage relative to expenditure data, we have no information about the value of different durable goods. Therefore, we can only use the population count of each durable good in the population and need to abstract from value weighted market size measure. The implication of this is discussed below in the section on matching.

[6]Cf. Beerli (2010) for a more detailed description of this data set.

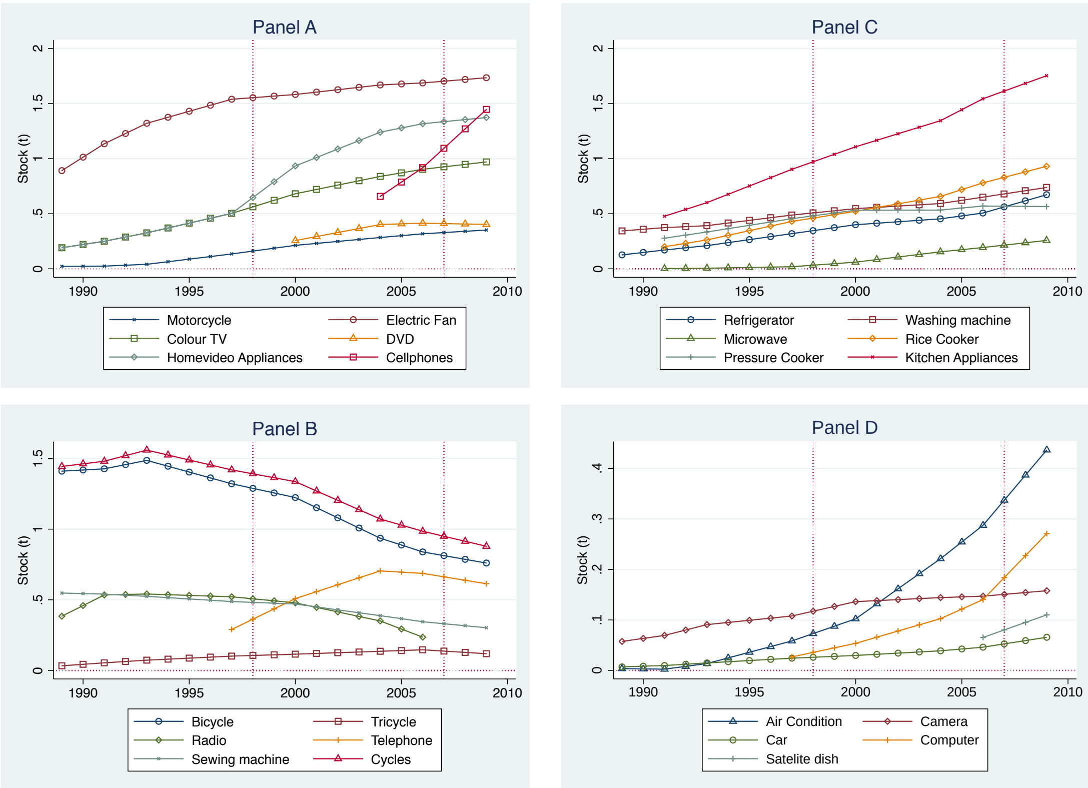
[7]As replacement rates of durable goods vary substantially across different households (cf. Bayus and Gupta, 1992) we abstract from including replacement demand as a constant fraction of each durable good's stock in our baseline

calculating the differences of durable good stocks at two different points in time, i.e.

$$Sales^{actual}_{j,t,t+k} = \left(Stock^{actual}_{j,t+k} - Stock^{actual}_{j,t}\right)\frac{1}{k} \tag{1}$$

The most natural way would be to look at durable good sales between two subsequent years which is interpreted as the flow of (new) durable good acquisitions adding to the stock, i.e. the market size of durable good j at time t.[8]

Figure 1: Evolution of Durable Good Stocks



Notes: CHNS data 1989 to 2009, missing years linearly interpolated. The figure shows for each durable good the number of items owned per household, i.e. $Stock^{actual}_{j,t}$, whereas "home video appliances" is the cumulative ownership of Colour TVs and DVD players, "kitchen appliances" the cumulative of microwave, pressure cooker and rice cooker and "cycles" the cumulative of bicycles and tricycles.

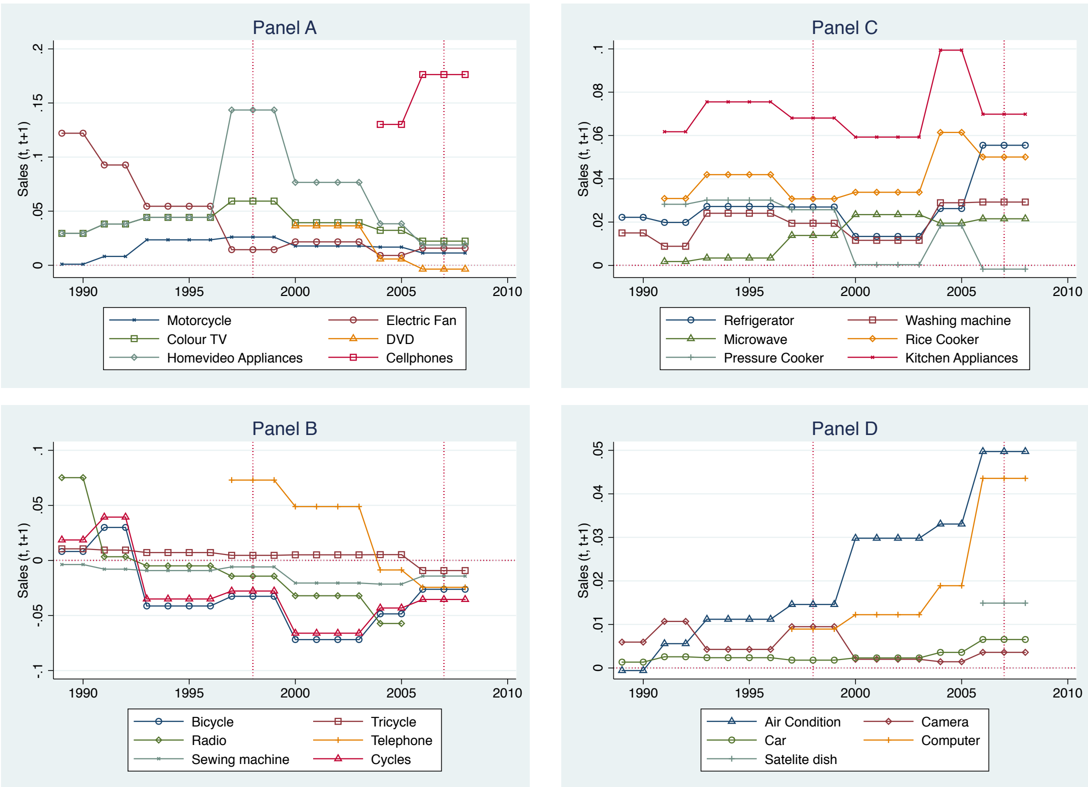
Figure 1 show evolution of the durable good stocks from 1989 to 2009, i.e. the diffusion of different durable goods. The time period from 1998 to 2007, when we observe firms' activities in our manufacturing data, is marked with the dotted vertical lines. It can be seen immediately that each

market size measure to avoid measurement error. This seems even more justified in the case of the largely unsaturated market in China.

[8]The last term, k, serves to annualise the flow measure and makes the effect of firms' demand expectation of different time horizons comparable. In the subsequent regression analysis we will look at a five year time horizon, i.e. we set $k = 4$.

good reveals its own characteristic diffusion pattern throughout the survey period.[9] The durable goods in panel A, for instance, diffused already in the early (electric fans) or late (colour TVs) 1990s, showing some signs of saturation in more recent years. An exception are cell phones which show rapid diffusion in the latest years when the CHNS began to cover it in the survey. Panel B depicts durable goods with decreasing ownership stocks in the latest years of the CHNS, such as bicycles, from which households substitute away to higher ranking goods such as motorcycles and cars. The durable goods in panel C and panel D show no signs of saturation yet, with the latter, such as cars or air condition appliances, taking-off in the latest survey years.

Figure 2: Evolution of Durable Good Sales



Notes: All CHNS data 1989 to 2009. The figure shows for each durable good, j the Sales variable $Sales^{actual}_{j,t,t+1} = Stock^{actual}_{j,t+1} - Stock^{actual}_{j,t}$. "Home video appliances" is the cumulative ownership of Colour TVs and DVD players, "kitchen appliances" the cumulative of microwave, pressure cooker and rice cooker and "cycles" the cumulative of bicycles and tricycles.

Figure 2 displays durable goods sales between two subsequent years throughout the survey period. Corresponding to the typical S-shaped diffusion pattern of durable good stocks, sales depict an

[9] As we will describe below, some goods could not be uniquely matched to manufacturing industries but are produced by the same four digit industries. We deal with this by creating a new ownership variable as the cumulative of all goods which belong to the same industry, i.e. “cycles” for bicycles and tricycles, “kitchen appliances” for rice and pressure cookers and microwaves and “home video appliances” for colour TV and DVD player ownership. Their figures are reported here for completeness.

inverted U-shaped pattern, increasing sharply when diffusion of the durable good in the population takes off, decreasing thereafter and leveling off when ownership approaches saturation. In addition, depending on the level of saturation, sales of different durable goods reach different maximum levels and need a different time interval for the same diffusion process. The figures show that even goods which are widely used in the population in most recent years, such as colour TVs and electric fans, reveal a very different sales pattern with fans decreasing throughout the survey period whereas colour TVs reach their peak in the late 1990s. Motorcycles parallel the hump shaped pattern of colour TVs across time but reach a much lower peak level. Most of the larger durable goods described above seem to be still in the increasing branch of the sales pattern in the latest survey years with refrigerators and washing machines showing non-monotonic pattern of sales and car sales slowly starting to take off.

Table 10 in the data appendix shows four year averages of the $Sales^{actual}_{j,t,t+1}$ variable over four consecutive periods covered by the CHNS. Table 11 in the data appendix shows the stock in each survey year. The tables give a more compact view on the dynamics discussed above and we can show with some back-of-the-envelope calculation what these numbers really mean. The stock of refrigerators, for instance, grew from 12% to 21% of all households between 1989 and 1993, a total increase of 8 percentage points or an average increase of about 2 percentage points per year. This means that in each year in the period between 1989 and 1993 an additional 2 percent of all households in China became new owners of refrigerators. How big are new refrigerators acquisitions equivalent to 2 percentage points of all Chinese households? With a very rough calculation we can take the total population in 1993 which was about 1'150 million people according to the World Bank (2012b), divide by the average household size of four[10] in the CHNS and take 2% of this number which yields an average yearly sales of 5.75 million refrigerators sold between 1989 and 1993.

In the next subsection we explain how we intend to relate the difference in the sales pattern we observe across different durable goods to innovation activities of firms and how household survey information and firm information are linked.

[10] Average household size was 4.14 people between 1989 and 1991 in the CHNS. This number is decreasing to 3.3 in the latest survey year 2009.

2.2 Manufacturing Data and Innovation Measures

Our main firm data set is the Annual Survey of Industrial Production (ASIP) from 1998 through 2007 conducted by the Chinese government's National Bureau of Statistics (NBS). The ASIP is a census of all non-state firms (not plants) with more than 5 million RMB in revenue (about $600,000) plus all state-owned firms in manufacturing. Manufacturing is defined here to include mining and public utilities. The raw data consists of over 150,000 firms in 1998 and grows to over 300'000 firms in 2007. The ASIP covers a wide range of information about the firm's balance sheet, cash-flow and on ownership which provides us with a rich set of control variables.[11]

The literature on endogenous growth theory outlined in the introduction gives no clear guidance on what kind of innovation activity applied researchers should focus when operationalising technical change. It simply says that R&D activity will be directed towards particular final goods or sectors which results in quality improvements or product inventions (i.e. new varieties). Those inventions, in turn, will increase the productivity of its firms or sectors relative to the others where no R&D activity was directed to. We make use of this open framework, employing different variables available in our data set as measures for all three stages in the R&D process: Innovation inputs, innovation outputs and productivity measures.[12]

We measure product innovation by the book value of new product sales, an innovation output variable directly available in the ASIP. According to the NBS (2012) new product are new with respect to the firm's prior product mix through improvements in the product's functionality and performance.[13] Although the definition entails some elements of process innovation by allowing products to differ from previous products through improvements in the production process, it is likely that quality improvements and process innovation are not fully captured by this variable (Jefferson et al., 2006). As an innovation input variable, we take investment in fixed assets as a proxy for R&D expenditure

[11] A detailed description of the data set can be found in Brandt et al. (2011).

[12] AL (2004: 1059) take also broad stand on the term "innovation", looking at drug approvals of the FDA as both kinds of outputs in the innovation process: Product innovation (what they call "new molecular entities") and process innovation respectively quality changes (what they call "generic drugs"). They justify looking also at quality improvements in the following way: "*Although generic drugs do not correspond to "innovations", their entry is driven by the same profit incentives as innovation*"

[13] The exact definition from the NBS (5.3.2012) (sent via email) is: "New product is the product which uses a new technological principle, new design, new concept, new production process, or has a marked improvement in structure, material and process with respect to the firms previous products, and thus significantly improves the product performance or expands the usability/functionality of the product.'

as the latter is available only for three years in our ASIP data. Investment in fixed assets has been widely used in the literature to proxy R&D expenditure as the adoption of new technologies and machines can be largely seen as a vehicle of technical change. The case is especially strong for a emerging market economy like China where technology adoption has a central role in technological progress.[14] In addition, employing physical investment has the the advantage that it is more widely reported and less skewed than R&D expenditure.

Furthermore, we look at the firm's productivity as a direct implication of successful innovation outcomes, both process and product innovation. Following the literature on firm-level innovation we focus on a simple measure of labour productivity, calculated as value added per worker.[15]

Table 13 in the data appendix shows the panel means of industries of our innovation variables. Interestingly, durable goods sectors are in general more innovative than the an average manufacturing firm with respect to each of our innovation measures. Furthermore, there are considerable level differences in innovation intensities across durable good sectors. In all innovation measures, car, computer and cell phone industry show relatively high innovation intensity levels whereas goods such as electric fans and cycles show low levels.

2.3 Matching Market Size Data and Innovation Data

The objective of this paper is to analyse how the changes in market size of different durable goods driven by the growing Chinese middle class affects the innovation activities of Chinese manufacturing firms. Now, which industries are *affected* by the changes in household demand?

In the CHNS we simply observe a household's ownership and change in ownership status of a specific durable good variety j and without having information on its price and quality. Dealing with such a population measure of market size has some implications.[16] First, we can not distinguish between car acquisition of one household to another household on a quality or price dimension[17]. All acquisition within the same durable good variety j receive the same (population) weight.[18] Thus, we think of

[14] Cf. Acemoglu et al. (2010). In fact, investment and R&D expenditure are highly correlated the in the three years of data we have at hand (correlation coefficient $\rho = 0.4$).

[15] Cf. the seminal contribution of Crépon at al. (1998) and Mairesse and Mohnen (2010) for a more recent review on applications of innovations surveys.

[16] Note that Acemoglu and Linn (2004) use a similar population measure of drugs used in a certain age group.

[17] This also includes second hand markets.

[18] Note that also acquisitions across time cannot be distinguished, although a car bought in 1989 and one bought in 2009 might, technically speaking, be very different durable good.

the new car acquisition, which we observe in the CHNS, as an average car bought or a count measure of sales whose magnitude can only be compared across durable goods. Second and related, we can not distinguish between sales values of similar magnitude between different durable goods. A 1 percentage point sale of cars and a 1 percentage point sale of bicycles affects their respective industries with a similar magnitude although an average car differs from an average bicycle to a large extent in value terms.

We link different durable goods observed in the CHNS simply to those four digit manufacturing sectors which produce them as a final household consumption goods by screening the NBS (2008) description of the Chinese Industry Classification (CIC) system. We neglect those manufacturing industries which produced the same durable goods but as equipment or intermediate inputs for other industries. Although including upstream and downstream industries of each final good sector into the analysis would give a richer picture of how demand effects the whole Chinese manufacturing sector, we restrict our analysis to final good sectors for the following reasons. First, investigations of the Chinese Input-Output table shows that the bulk of value added is generated within each final good sector and less so in related industries. Thus, if we cannot find an effect in the final good industries, up- and downstream sectors are even less likely to respond. Second, matching the CHNS population market size measure to value added figures in the Input-Output tables involves taking many additional, implicit assumptions with respect to relative good values which would increase measurement error in our market size measure to a large extent.

We arrive at a total sample of 16 different manufacturing industries, starting with an initial set of 22 different durable goods.[19] Since colour TVs and DVD players are produced by the same four-digit manufacturing industries, we created a new ownership for home video appliances simply as the cumulative of those two goods irrespectively whether this is a colour TV or a DVD player. Similarly, we created variable for kitchen appliances for microwaves, rice cookers and pressure cookers which fell all into household kitchen appliances manufacturing industry.[20] Cell phones and telephones also belong to the same manufacturing industry in the SIC before 2002 but are two different industries thereafter.[21] Since cell phone ownership is only available in the CHNS from 2004 onward but tele-

[19]The exact list of durable goods and matched industries can be found in Table 14 in the data appendix.

[20]Colour TVs and DVD players are both produced by the industry "Home video equipment manufacturing" (CIC 4071). Rice cookers, pressure cookers and micro waves are all produced by the industry "Household kitchen appliances manufacturing" (CIC 3954) and bicycles and tricycles are produced by the industry "bicycle manufacturing" (CIC 3741).

[21]In 2003, the CIC system was revised to include more detail for some sectors, while some others were merged. Part

phone ownership is available throughout the sample. We follow Brandt et al. (2011)[22] and exclude firms which are not in manufacturing sectors, firms with less then 8 employees and with negative values of value added and capital stock. We end up with a final sample of 1'925'846 manufacturing firms and 34'324 durable good firms in the years 1998 to 2007.

2.4 Potential Market Size and Identification

A major difficulty in any investigation of the impact of market size on innovation is the endogeneity of actual market size, i.e. better and cheaper products having larger markets. We follow Acemoglu and Linn (2004) in constructing a *potential market size* measure where the acquisition of new durable goods is only driven by changes in the income distribution and remains unaffected by changes in quality or prices of durable goods. We construct our potential market size measure exploiting the quasi-exogenous[23] component of market size driven by dynamics in the income income distribution, combined with differences in the ownership profile across income groups for different kinds of durable goods. We obtaining ownership profiles of different income groups and changes in the population shares of those income groups from the CHNS. Following Acemoglu and Linn (2004) our potential stock measure and, correspondingly, our potential market size measure of sales of for durable good j at time t is

$$Stock_{j,t}^{potential} = \sum_g \bar{u}_{j,g} i_{j,t} \tag{2}$$

$$Sales_{j,t,t+k}^{potential} = \left[Stock_{j,t+k}^{potential} - Stock_{j,t}^{potential}\right] \frac{1}{k} = \left[\sum_g \bar{u}_{j,g}\left(i_{g,t+k} - i_{g,t}\right)\right] \frac{1}{k} \tag{3}$$

where $i_{g,t}$ is the population share of income group g at time t in our CHNS sample and $\bar{u}_{j,g}$ is the usage profile of durable good j in income group g in the base-year 2009, i.e. $\bar{u}_{j,g} = u_{j,g,t=2009}$. It is important to that over-time variation in this measure is not from changes in individual use, but results completely from changes in the income distribution, i.e. through changes in the size of income groups $i_{g,t}$. Consequently, changes in price and durable good quality, which may result from innovation and affect consumption patterns, will not cause over-time variation in $Sales_{j,t,t+k}^{potential}$.

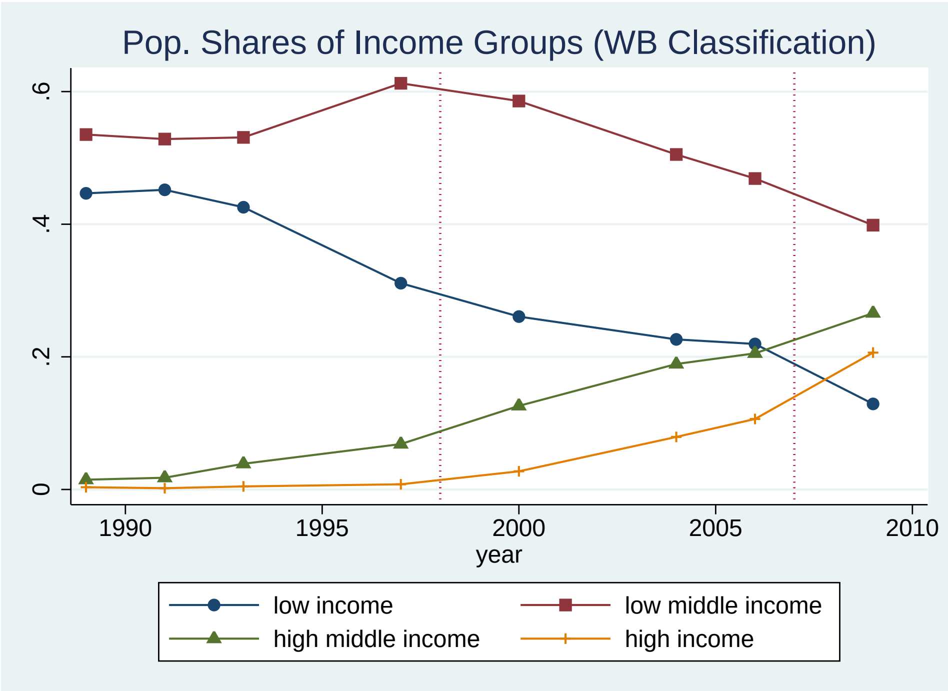
of this is splitting up the "Mobile Communication Terminal Equipment" industry from the "Communication Terminal Equipment". See Table 14 in the data appendix for industry codes.

[22] We also employ their procedure to link restructured firms over time, cf. the online appendix of Brandt et al. (2011) for more details.

[23] We discuss in the section on the empirical strategy why we think reverse causality, i.e. the effect of innovation activities on dynamics of the income distribution, is not a first-order concern for our identification strategy.

We compute $i_{g,t}$ by splitting the income distribution into G groups, i.e. $g = 1, ..., G$, by setting fixed income thresholds in constant 2009 Yuan. In our baseline measure we make use of a categorisation of the World Bank (WB) (2009), which leaves us with four groups (low income, lower middle, upper middle, high income).[24]

Figure 3: Evolution of Income Groups According to WB Classification



Notes: All CHNS data 1989 to 2009. Households classified into four income groups according to their household income per capita in constant 2009 Yuan: low income (2'150 Yuan), lower middle income (2'150 - 8'515 Yuan), upper middle income (8'515 - 16'500 Yuan), high income (16'500 or more). The group "total low" shows the combination of low and low middle income group, whereas the group "total middle" shows the combination of low middle and high middle income group.

Figure 3 shows the evolution of population shares of the four income groups across the survey period. According to this definition of income groups, the population share of poor (low income and lower middle) fell dramatically, especially from 1997 (still 98 %) to 2009 (around 52%). Conversely, the total size of middle class (and rich) members increased rapidly in the second half of the 1990s reaching 87% of the population in 2009. Interestingly, the share of highest income group with considerable

[24]Households were assigned to four income groups according to their household income per capita following a classification of the WB atlas method that assigns economies into 4 groups according to their GNI per capita in 2009 (cf. WB (2012a) for a more recent description). The groups are: low income, \$1'005 or less; lower middle income, \$1'005 - \$3'975; upper middle income, \$3'975 - \$7'675; and high income, \$7'675 or more. The threshold for the high income group was substantially lowered in order to get an accurate measure of the usage intensities also in early years of the CHNS (when few households are rich). All dollar figures where converted into constant 2009 Yuan using the exchange rate and PPP adjustment factors. Cf. notes of figure for thresholds in Yuan.

more purchasing power grew especially strong in the latest survey decade with the high income group increasing its share from a mere 1% in 1997 to 20% in 2009.

We calculate the usage profile, $\bar{u}_{j,g}$, for each durable good j of each income group group g as the number of items per household in that income group in the base-year 2009, i.e. the group's stock of durable j. Choosing among different CHNS waves as base-year implies different assumptions about entrepreneurs' expectations, on the one hand, and raises issues of data availability and accuracy on the other hand. Because the 2009 wave of the CHNS has the richest coverage of durable goods and the highest income group is sampled more accurately than in earlier years, we pick 2009 as our best choice of a base-year. For a more detailed discussion of this issue and robustness considerations we refer to the data appendix.

Table 12 in the data appendix shows the usage profiles for the four income groups with the base-year 2009. We can compare the usage intensities, $u_{j,g,t=2009}$, of adjacent income groups to gauge the effect of households moving up the income ladder into higher income groups. For this reason we report the differences of the usage intensities for adjacent income groups in the second row of each durable good in brackets.
Imagine a hypothetical household, moving from the lowest income group to the low middle income group. The average number of electric fans used by this household would increase from 1.43 to 1.78 (an increase of 0.35), whereas moving from the lower middle to the higher middle income group would only result in an increase of 0.01 (to a stock of 1.79). On the other hand, refrigerators experience the highest relative increase in stock (0.25) if the household would move from the lower middle (stock: 0.54) to the higher middle income group (0.79) whereas moving to the richest group (stock: 0.92) or moving out of the poor group (stock: 0.45) does not results in higher changes. For cars, the increase is highest if a household would move out of the upper middle (stock: 0.06) into the highest income group (0.13).

How do the characteristics in these usage profiles relate to dynamics of sales of different durable goods, on the one hand, and to changes in income groups, on the other hand? We observed above that there was a massive reduction in poverty with the share of the low income group dwindling already in the early 1990s whereas the lower middle income group's share first increased in the 1990s and then started to fall monotonically thereafter. Conversely, the share of the upper middle income group started to increase with large steps after 1997 whereas the high income group experiences a more moderate growth since 2004.

As one would expect, the acquisitions of durable goods parallels the dynamics in the income distributions (cf. Figure 2). Clearly, most lower ranking goods such as colour TVs, electric fans and motorcycles experienced a typical hump-shaped sales pattern with a peak in the 1990s with their sales declining thereafter. On the other hand, the durable goods with the largest increase in ownership in the highest income group, such as cars and air conditions, still experience upward shifts in sales in 2009. The durable goods which experience the highest increase in usage in the middle income groups show a somewhat intermediate pattern of sales in more recent survey years.[25]

Next, we would like to analyse the evolution of different innovation measures across industries. In Figure 4 different industries are allocated to three groups with different (expected) dynamics of their market size from 1998 to 2007, the period covered by our ASIP data. According to the evolution of their sales in this time window and depending on which income group shows the highest increase in usage intensity in the base-year, we allocate industries belonging either to "take-off industries", "post-take-off industries" and "saturation industries".[26] Although the figures below use only a small fraction of the information we will use in the regression analysis, it serves as a good illustration and provides suggestive evidence for our basic story.[27]

Panel A of Figure 4 plots the evolution of each groups' share of new product innovators across time,

[25]One interesting feature is the difference in the usage of cell phones and computers across income groups and the corresponding difference in diffusion. Both goods can be considered as "new inventions" and were introduced in the survey relatively late, computers in 1997 and cell phones in 2004. Differently to computers, cell phones are widely used already in poor income groups and seem to experience only modest increases when moving to higher income groups. Computers, however, are quite differently used across income groups, with the largest increase in usage being between the upper middle and the high income group. This pattern translates to the observed diffusion rates and sales rates; Cell phone stock and sales sky-rocked since their introduction with everybody, also the poor households, being able to acquire an item whereas computer stocks and sales are increasing too but by a more modest extent.

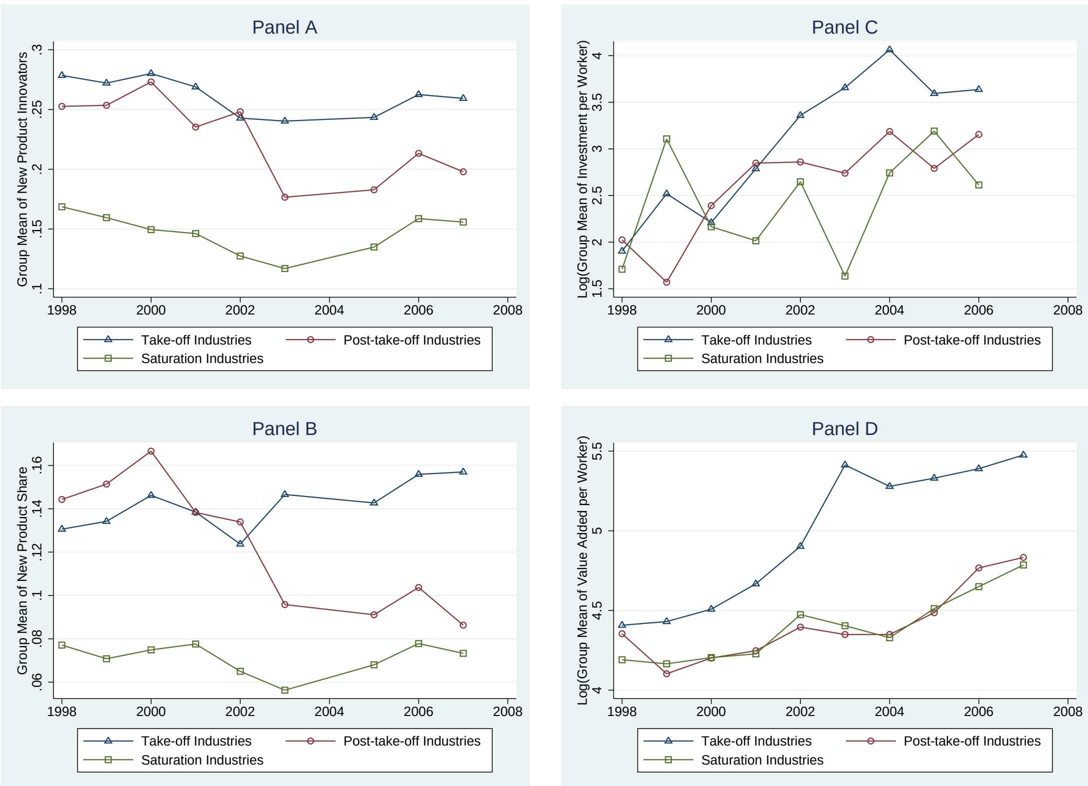
[26]We allocated all durable goods with the highest increase in their ownership intensity in the high income group (according to Table 12 in the data appendix) into the group "take-off industries", i.e. air condition, camera, cars, computers and kitchen appliances. Durable goods with the highest increase in their usage profiles moving from the poor in to the lower middle income group were allocated into "saturation industries" (i.e. cycles, electric fans, home video appliances, motorcycles) plus all other durables with declining stock, i.e. telephones, sewing machines and radios. Kitchen appliances, refrigerators and washing machines with their highest increase in the usage intensity in the middle income group were allocated to "post-take-off industries". Although the cellphone industry experiences the highest increase in its usage intensity when moving from the low income group to the low middle income group, its sales sky-rocked in the last years of our observation period when it was introduced in the CHNS. Therefore, we allocate it also to the "take-off industries". We treat the satellite dish industry similarly.

[27]We are particularly aware of the fact that the level differences in the innovation variables reported in Table 13 in the data appendix may drive the following results to some extent.

our innovation output measure. The "take-off industries" emerge as the group with the highest overall number of innovators (26 percent of all firms on average) whereas "post-take-off industries" and "saturation industries" have clearly a lower share. Although the share of innovators is decreasing in all groups, "take-off industries" experience the lowest decrease. Panel B of Figure 4 plots the average share of new product sales (on total sales) for each group, showing that the average share of new product sales increased by 20 percent in "take-off industries" from 1998 to 2007 whereas it decreased for the other two groups.

Looking at our innovation input measure (cf. Panel C of Figure 4) reveals that investment per worker increased to a large extent (466 percent) in "take-off industries" whereas the increase was much lower for the other two groups, i.e. 210 percent for "post-take-off industries" and 147 percent for "saturation industries". Innovation performance shows a similar picture (cf. Panel D of Figure 4). Average labour productivity increased by almost 200 percent in "take-off industries", whereas growth was clearly less pronounced in the other two groups.

Figure 4: Evolution of Innovation Variables in Industry Groups



Notes: All ASIP data 1998 to 2007. Industries allocated to groups according to the evolution of their sales from 1998 to 2007 and depending on which income group shows the highest increase in usage intensity in base-year. Take-off industries are air conditions, cameras, cars, cellphones, computers and kitchen appliances. Saturation industries are cycles, electric fans, home video appliances, motorcycles plus other durables with declining stock, i.e. telephones, sewing machines and radios. Refrigerators and washing machines were allocated to post-take-off industries.

2.5 Empirical Strategy

We consider the following baseline regression model

$$Inno_{i,j,t} = \alpha_1 Sales^{actual}_{j,t,t+4} + \mathbf{X}'_{i,j,t}\alpha + \psi HHI_{j,t} + \eta_j + \lambda_t + \epsilon_{i,j,t}, \tag{4}$$

where i denotes a firm, j a sector and t the time. In our baseline specification, our independent variable are actual sales of durable goods, where $Sales^{actual}_{j,t,t+4}$ is defined as the average annualized change in actual ownership rates between t and $t+4$. Technically, our data would also allow us to look at shorter time horizons or slightly longer time horizons. For comparability reasons, however, we follow Acemoglu and Linn (2004) and look at a five year future time horizon assuming that this is the relevant time period for firms to decide on their innovation activities.[28]

The rich dataset gives us the possibility to analyse the market size effect on different dimensions of firms' innovation activities. In the first baseline specification, the dependent variable, $Inno_{i,j,t}$, measures labour productivity. Labour productivity is defined as the value added per worker and is a direct outcome of innovation performance. Further, we use investment as a measure for innovation input and a dummy for new product sales to measure product innovation output.[29] We measure all dependent variables on the firm level and estimate the first two in a standard OLS framework and the later with a Linear Probability Model.

All specifications include sector fixed effects, η_j, to account for sector specific differences in innovation opportunities and time fixed effects, λ_t, to account for aggregate trends. The vector $\mathbf{X}_{i,j,t}$ controls for unobserved firm-level heterogeneity including the logarithm of firm size (employment), age, dummies for ownership and whether a firm is located in a coastal province. Moreover, since literature has pointed out the market concentration as important determinant of firms' innovation activities, we introduce a measure of market concentration, $HHI_{j,t}$, in our baseline regression.[30] The Tables 15 and 16 in the data appendix show descriptive statistics for the set of firm controls we employ in our analysis.

The coefficient of interest, α_1, captures the effect of market size on a firm's innovation activity and we expect it to be positively associated with market size. In order for this coefficient to be estimated

[28] Changing the time horizon leaves our results qualitatively unchanged, although reducing the precision of our estimates in some cases.

[29] We are aware of the fact that investments are an imperfect measure for innovation inputs. However, due to high correlation we regard it as the first best proxy to measure R&D investments. Cf. Section 2.2 for a discussion.

[30] For instance, Aghion et al. (2005) find evidence for an inverted-U relation between competition and innovation within industries in a study on a panel of British firms.

consistently, we need to address several concerns (cf. Section 3.2 and Section 3.3). First, firms might take not only the domestic market into consideration when implementing R&D, especially in China with its export-led growth. We control for this issue by including a export dummy indicating whether exporting firms systematically differ from non-exporting firms. We will further include an interaction term between market size and the intensity of export behaviour (as measure by the sales that accrues due to exports) to get an idea about the effect of the intensive margin of exports.
Second, a major threat to the validity of our empirical strategy are time-varying omitted variables, particularly supply side technology shocks, since sector fixed effects account only for non-time varying differences in technology opportunities across sectors. To control for time-varying technology opportunities we supplement our baseline regression specification including a measure of worldwide technology potential reported by Swiss firms in several years on a highly disaggregate sector level.[31]
A third and important issue, is the endogeneity of actual market size discussed at length above. As we measure actual market size in the household data and on the sector level, the possible reverse causality link of innovation activities of individual firms affecting market size on the sector level is not completely obvious. Still, even on the sector level, we think the potential endogeneity of market size is an important issue and we employ an instrumental variables strategy fleshed out in full detail in Section 3.3.

[31]The KOF Innovation Survey (KOF, 2012) covers a representative sample of Swiss firms in the manufacturing, construction and service sector on a three yearly basis since 1990. Firms assess on a five-point Likert scale the "technology potential", i.e. the world wide availability of technological know-how in private and public hands which could be used to generate marketable new products. To the best of our knowledge, the KOF Innovation Survey is the only publicly available innovation survey which can be used on a highly disaggregate sector level (four digits). We match the "technology potential" as assessed by Swiss firms to ASIP firms using sectors and time. Additionally, we check for robustness of this measure using standard innovation measures such as R&D spending, the number of patents and new product outputs share on the same sector level.

3 Results

3.1 Baseline Specification

Market Size Effect on Labour Productivity

We start out, estimating a version of our baseline regression equation 4 with sector and year fixed effects and evaluating the effect of switching in the different firm level controls. As outlined above, our dependent variable is given by the firm specific labour productivity (measured in logs). Although we started our analysis with 16 durable goods, we will drop the two goods "Cellphones" and "Satellite Dishes" due to an insufficient number of observations for the following estimation approach.[32] Moreover, since our market size measure is forward looking by four periods, we lose the last two recent years from our firm level panel. Hence, we are left with a panel of 14 (durable good) sectors and 8 years (1998-2005).

Column (1) of Table 1 shows a highly significant effect of our actual market size measure on labour productivity. Increasing the average annualized market size by one percentage point raises firms' labour productivity by 2.48%.[33] To get a better feeling for the magnitude of the effect, let us take a closer look at the dependent variable. In fact, we find the logarithm of labour productivity having a mean of 3.71 with a standard deviation of 1.32. Hence, an increase in market size by one percentage point translates into an increase in labour productivity by 1.88 standard deviations.[34]

Although the industry fixed effect controls for fundamental (non-time-varying) differences between sectors, firm composition of sectors could be very different explaining large parts of the observed differences in innovation activities across sectors. To make sure that our estimates are not biased by omitted variables which are frequently discussed as firm-level determinants of innovation activities, we exploit a large set of available information such as firm size, age, ownership, location and market concentration:[35] As suggested in the literature we use the log of workers as proxy for firm size and use a dummy for firms that are older than six years (the median in our sample).[36] We include a dummy for firms located in coastal provinces, worrying that firms in the booming coastal regions

[32]For the two goods of "Cellphones" and "Satellite Dishes", information on ownership becomes available only in 2004, 2006 respectively.

[33]Remember that we allow households to own more than one item of each durable good.

[34]The standardized coefficient is derived by dividing the estimation coefficient in Table 1 with the regressand's standard deviation.

[35]For detail, the reader is referred to Crépon et al. (1998) and Mairesse and Mohnen (2010) for a review of firm-level innovation determinants.

[36]Arnold and Hussinger (2005) for example argue that due to possible correlation between size and age of a firm employing a dummy instead of the absolute age seems to be the correct estimation approach.

Table 1: OLS Regression of Labour Productivity

Dependent Variable: Labour Productivity

	(1)	(2)	(3)	(4)
$Sales^{actual}_{j,t,t+4}$	2.484**	3.144***	3.144***	3.144**
	(0.994)	(0.924)	(0.691)	(1.248)
Size		0.00422	0.00422	0.00422
		(0.0124)	(0.0164)	(0.0358)
admin_FE		0.322***	0.322***	0.322***
		(0.0316)	(0.0293)	(0.0469)
admin_SOE		-1.063***	-1.063***	-1.063***
		(0.0610)	(0.0452)	(0.0795)
admin_COE		0.0134	0.0134	0.0134
		(0.0287)	(0.0274)	(0.0517)
Age		-0.150***	-0.150***	-0.150***
		(0.0225)	(0.0246)	(0.0385)
Coastal		0.0960**	0.0960**	0.0960
		(0.0427)	(0.0367)	(0.0811)
HHI		2.31e-05	2.31e-05	2.31e-05
		(2.71e-05)	(2.62e-05)	(3.95e-05)
Clustering	Firm	Firm	Sector×Year	Sector
No of Clusters	8112	8110	109	14
Observations	21,853	21,843	21,843	21,843
R-squared	0.081	0.175	0.175	0.175

Notes: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent. Clustered standard errors in brackets. All regressions include year fixed effects and industry fixed effects. $Sales_{t,t+4}$ is defined as the average annualized change in ownership rates between t and t+4. Labour productivity is calculated as the log of value added per worker.

might be overrepresented in some sectors. Additionally, we control for the ownership structure of firms as it was shown in the literature that strong ties to the state might influence firms financing and productivity.[37] Specifically, we take privately owned firms as the reference group and introduce three dummy variables, $admin\_FE$, $admin\_SOE$ and $admin\_COE$ to indicate whether a firm is foreign-owned, state-owned or collectively-owned respectively. Finally, to control for different intensities of market competition across sectors, we introduce the Hirschmann-Herfindahl index, which is defined as the sum of squared market shares of all firms within the sector.[38]

Column (2) in Table 1 shows the more demanding specification with the full set of firm-level controls. Controlling for firm heterogeneity increases the market size effect by almost one third. Statistically significant at the one percent level, one percentage point increase in market size leads to an increase in sector specific labour productivity by 3.14%. Moreover, the R^2 statistic increases indicating a much better fit of the model. As expected, the foreign-owned firms and those located at the coast seem to be more productive while state-ownership and age decreases a firm's labour productivity. Size and market concentration does not have an effect on labour productivity.

Since we look at firm level outcomes but our measure of market size varies at the sector level, accounting for correlation of regression residuals within sectors is potentially important.[39] While the standard errors in Columns (1) and (2) were clustered on the firm level, Columns (3) of Table 1 reports standard errors clustered on the sector-time dimension. Column (4) reports the most conservative approach, which uses sector-level clustered standard errors. However, the cluster estimator is only consistent as the number of clusters gets large and, in our case, the number of available sectors is quite small (fourteen), it is very likely that standard error clustering on the sector level does induce further precision problems. As we cannot increase the number of available durable goods (and hence clusters), we pick an intermediate level and cluster by sector and year as our baseline and report, as proposed by Angrist and Pischke (2009), estimates for sector averages in Table 7. With this approach we allow for within sector and within time correlation among residuals, and we get a

[37] See for example Song et al. (2011) and Brandt et al. (2011).

[38] Studies that specifically employ the HHI are for example Cotterill (1986), Farrell and Shapiro (1990) and Farrell and Shapiro (1990). We define the HHI for industry j at time t as the sum of squared market shares (in value added) of all firms operating within this sector at time t. Since we calculate market shares in percentage terms, (between 0 and 100), the HHI lies in the range between 0 and 10 000. We are aware of the fact that the border of markets is less clear for globally operating firms. However, we consider the HHI as the first best measure to capture market competition within the firm's primary (home) market.

[39] Note that the sector fixed effect, η_j, removes the sector mean, $\bar{\epsilon}_j$, from the error term, $\epsilon_{i,j,t}$. Nevertheless the residual errors, $\epsilon_{i,j,t} - \bar{\epsilon}_j$, might still be correlated within sectors and, serially, across time.

reasonable number of clusters (109 as seen in the last row of Table 1). We see that the estimate of a positive market size effect is robust to different levels of clustering and even becomes more precisely estimated in Column (3) of Table 1.

Market Size Effect on Innovation Inputs and New Product Innovation

Table 2 displays the estimation results for log investment as our dependent variable. The results indicate significant evidence for a positive effect of actual market size on firms' innovation performance. For instance, Column (1) of Table 2 shows that an increase in the annualized average market size by one percentage point increases firm's innovation inputs by 6.31%. Adding firm control variables reduces the estimate in size, but it remains strongly significant. These findings are robust to introducing the more conservative level of clustering in Column (3). For completeness, we again report the results for sector level clustering in Column (4) of Table 2 even though the number of clusters is insufficient again.

Table 2: OLS Regression of Log Investment

Dependent Variable: Log Investment

	(1)	(2)	(3)	(4)
$Sales^{actual}_{j,t,t+4}$	6.310***	4.065**	4.065***	4.065*
	(2.185)	(1.749)	(1.242)	(2.282)
Firm Controls	No	Yes	Yes	Yes
Clustering	Firm	Firm	Sector×Year	Sector
No of Clusters	5680	5680	109	14
Observations	12,832	12,829	12,829	12,829
R-squared	0.063	0.384	0.384	0.384

Notes: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent. Clustered standard errors in brackets. All regressions include year fixed effects and industry fixed effects. Columns 2-4 include firm controls of size, ownership, age, location and market competition. $Sales_{t,t+4}$ is defined as the average annualized change in ownership rates between t and $t+4$.

To complete our picture on firm's innovation behaviour in the advent of a rise in demand, we explore the effect of an increasing market size on product innovation on the firm level. In contrast to labour productivity which results from both product and process innovation, reported new product sales measures directly the success of a firm's product innovation activities. We use a dummy for whether a firm has positive new product sales as a relative robust measure of product invention.[40] To estimate

[40]Employing a dummy for positive new product sales has some robustness advantages and simplifies the regression

the market size effect for this dependent variable, we use the linear probability model which allows us to interpret the estimated coefficient directly as the average probability of a new product innovation.[41] Table 3 displays the estimation results with actual market size over a four year time window. The baseline estimation in Column (1) of Table 3 shows a significant positive effect of actual market size on a firm's probability to innovate new products. In column (2) we introduce again controls on the firm level (in addition to the fixed effects) to test the robustness of this estimate. The estimated coefficient of market size reduces slightly but remains significant at the five percent level. Column (3) corrects the standard errors for possible correlation within sectors and time and Column (4) shows the most conservative approach of sector clustered standard errors again. For instance, Column (3) in Table 3 shows that an increase in market size by ten percentage points increases the probability of a successful product invention by 7.43%.

Interestingly, the effect of the specific firm characteristics (captured by the covariates) seem to be very different for product innovation. In contrast to the two other innovation variables, the most efficient firms in terms of product innovation are state-owned enterprises that are large in size, old and located in inland-provinces. Robust to all different specifications, foreign ownership seems to have a negative impact of the probability to invent new products.[42]

3.2 Robustness Checks

Open Economy Consideration

Our baseline estimations displayed in the previous section assumed that China was a closed economy with firms operating only in the domestic market. As the Table 16 in the data appendix shows, this is clearly not a very realistic assumption, as roughly 50 percent of all firms in durable good industries do export with their average export share of total sales being roughly 30 percent. Especially if the majority of output within one sector is sold abroad it should be the world demand structure that

---

analysis (in fact only roughly 20 percent of firms are inventors in our sample as evident form Table 13), on the one hand, but we lose information on the intensive margin of product innovation, on the other hand.

[41] An alternative estimation approach is to use a discrete choice model. Hence, we re-estimated the effect using a probit and a logit model, and the results remained very similar.

[42] In fact, 36 percent of all state-owned firms are new product inventors whereas this number is considerable smaller for private owned firms (11 percent) and for foreign owned firms (17 percent). At the first glance puzzling, we interpret this finding as a consequence of inferior knowledge of local market conditions and consumers' preferences by foreign-owned firms. As the introduction of new products always poses a risk to firms (not knowing the ex-post realized demand for it), Chinese-owned enterprises should have an advantage in estimating Chinese preferences such that their propensity to innovate new products is higher.

Table 3: Regression of Positive New Product Sales

Dependent Variable: Dummy for Positive New Product Sales				
	(1)	(2)	(3)	(4)
$Sales^{actual}_{j,t,t+4}$	0.888***	0.743**	0.743***	0.743*
	(0.305)	(0.293)	(0.236)	(0.389)
Firm Controls	No	Yes	Yes	Yes
Clustering	Firm	Firm	Sector×Year	Sector
No. of Clusters	7554	7552	96	14
Observations	18,323	18,313	18,313	18,313

Notes: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent. Clustered standard errors in brackets. All regressions include year fixed effects and industry fixed effects. Columns 2-4 include firm controls of size, ownership, age, location and market competition. The dependent variable is a dummy which equals one if a firm has positive sales due to new products and equals zero otherwise. $Sales_{t,t+4}$ is defined as the average annualized change in ownership rates between t and $t+4$.

determines the sector's market size and consequently the firm's innovation activities.[43] If world demand is correlated with our measure of market size, our previous estimate would be biased and would not capture the effect of domestic demand on firm's innovation activities. Moreover, exporting firms may be different per se from firms serving only the domestic market. The economic literature has long stressed the interference between openness, free trade and economic growth.[44]

It is, thus, important to control for exports and we include a dummy for whether a firm has positive exports, $EXP_{i,j,t}$, in our baseline regression framework:

$$Inno_{i,j,t} = \alpha Sales^{actual}_{j,t,t+4} + \mathbf{X}'_{i,j,t}\beta + \psi HHI_{j,t} + \gamma EXP_{i,j,t} + \eta_j + \lambda_t + \epsilon_{i,j,t}. \tag{5}$$

Column (2) of Tables 4, 5 and 6 show the estimation output for all three dependent variables. Column (1) in each table repeats the baseline estimation from the previous section for comparability, where the baseline specification includes all firm controls of size, ownership, age, region and market concentration and standard errors are clustered by sector and year (see Column (3) of Tables 1, 2 and 3). In line with the existing literature the coefficient on exports is positive and strongly significant in all specifications, i.e. exporting firms seem to be more involved in innovation activities than non-exporting firms. Across all specifications the baseline results stay qualitatively and quantitatively

[43]Note that only the radio and camera industry experience export shares higher then 50 percent of their total sales.

[44]Even if the direction of causality is less clear, the pure existence of correlation between both variables is not questioned. Ben-David (1993) finds significant correlation between trade liberalization and income convergence on the country level, while Arnold and Hussinger (2005) study the relation between export intensity and productivity growth within a German firm panel data set.

very similar and the market size effect becomes even slightly stronger.[45] For instance, Table 4 shows that an increase in market size by one percentage point increase labor productivity by 3.40% instead of 3.14% as soon as we control for export behaviour of firms.

Now, the specification above cannot distinguish between firms exporting only a small share of their sales and firms serving almost exclusively the external market. To get an idea of the effect on the intensive margin of export behaviour, Column (3) uses a slightly different specification including an interaction term between market size and the firm's exportshare (as a share of total sales).[46] Column (3) of Tables 4, 5 and 6 show the estimation output for all measure of innovation performance. Interestingly, we find for all specifications that higher exportshares are negatively associated with firms' innovation activities. Moreover, the interaction term between our variable of interest, industry specific market size, and firms exportshares is strongly negative and significant. These findings suggest several things: first, exporting firms are on average more innovative than their counterfactual only serving the local market. Second however, the innovation activities decrease in the intensive margin of exports and third the market size effect on innovation activities is smaller for firms that export more of their total sales. These findings are in line with our expectations as for firms that almost exclusively sell offshore should be concerned about the global demand rather than the local demand.

Supply Side Determinants of Innovation

A considerable threat to the validity of our empirical strategy is that firms' innovation decisions are mainly determined by supply side factors. It is straight forward to conjecture that sectors have very different innovation profiles, e.g. with the car industry investing a lot more in R&D than the much smaller sewing-machine industry. Non time-varying differences in technological opportunities across sectors are accounted for by the sector dummies. However, technology opportunities might change over time and they might affect different sectors (the time dummy controls for the aggregate trend). We tackle this issue by including a measure of worldwide technology potential on the sectoral level in our regression. As described in the empirical strategy, this variable captures "technological opportunities" as assessed by Swiss firms in to whole set of manufacturing sectors on a five point

[45] Moreover, the effect of other firm-level covariates does not change either.

[46] The regression specification becomes $Inno_{i,j,t} = \alpha Sales^{actual}_{j,t,t+4} + \mathbf{X}'_{i,j,t}\beta + \psi HHI_{j,t} + \phi \left\{Sales^{actual}_{j,t,t+4} \cdot EXSH_{i,j,t}\right\} + \gamma EXSH_{i,j,t} + \eta_j + \lambda_t + \epsilon_{i,j,t}$ where $EXSH_{i,j,t}$ denotes the firm i specific share of total sales accruing to exports at time t in sector j.

Likert-scale (KOF, 2012).[47] As Table 17 in the data appendix shows, the variable shows considerable variation across time and sectors. We estimate the following regression model:

$$Inno_{i,j,t} = \alpha Sales^{actual}_{j,t,t+4} + \mathbf{X}'_{i,j,t}\beta + \psi HHI_{j,t} + \gamma EXP_{i,j,t} + \delta Inno^{Swiss}_{j,t} + \eta_j + \lambda_t + \epsilon_{i,j,t}, \qquad (6)$$

where $Inno^{Swiss}_{j,t}$ indicates the worldwide technology potential on sectoral level reported by Swiss firms. The last Column of Tables 4, 5 and 6 show the estimated coefficients that consistently verify the validity of our previous estimates. Using labour productivity as dependent variable, Table 4 shows that the coefficient on market size increases compared to the previous specification in Column (2) and is strongly significant at the one percent level. Looking at innovation inputs and the probability of a new product invention does not alter the results. As seen in Tables 5 and 6 introducing the control variable of sector specific technology potential leaves the coefficient on the market size effect essentially unchanged and highly significant at the one percent level. Now, what is interesting to observe is the effect of the technology variable itself: While the coefficient is insignificant for a firms decision of innovation inputs and new product output, it turns negative for labour productivity (see Column (4) in Table 4.

Although this finding is surprising on a first glance, different explanations are possible: first, bearing in mind that the variable captures the potential of the most advanced technologies but China is still an emerging market, Chinese firms may not be affected by changes at the world frontier technologies. Instead, innovation activities of Chinese firms might be of a more incremental nature, developing technologies aimed at the local market. Second and more important, our findings actually can be read as a conformation to the "Hypothesis of unsuitable technologies" which states that countries/sectors that use the most advanced technologies but lack the required human capital to operate them may be worse off and not use them (see e.g. Acemoglu and Zilibotti (2001)). Overall these findings make us confident that the demand-induced market size effect is not driven by fundamentally different technological opportunities across sectors.

[47] We aggregate the firm-level values of this variable to sector means and match it to the ASIP using sector and time identifiers.

Table 4: Robustness Check: Labour Productivity

Dependent variable: Labour Productivity

	(1)	(2)	(3)	(4)
$Sales^{actual}_{j,t,t+4}$	3.144***	3.401***	4.065***	3.657***
	(0.691)	(0.745)	(0.666)	(0.728)
Sales_Exportsh			-3.036***	
			(0.794)	
Exportshare			-0.265***	
			(0.0501)	
Exporting		0.135***		0.135***
		(0.0312)		(0.0313)
Techn_Pot				-0.0431*
				(0.0230)
Firm Controls	Yes	Yes	Yes	Yes
Clustering	Firm	Firm	Sector×Year	Sector
No of Clusters	109	109	109	109
Observations	21,843	21,770	21,770	21,770
R-squared	0.175	0.170	0.174	0.170

Notes: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent. Clustered standard errors in brackets. All regressions include year fixed effects, industry fixed effects and firm controls of size, ownership, age, location and market competition. $Sales_{j,t,t+4}$ is defined as the average annualized change in ownership rates between t and $t + 4$.

Table 5: Robustness Check: Log Investment

Dependent Variable: Log Investment

	(1)	(2)	(3)	(4)
$Sales^{actual}_{j,t,t+4}$	4.065***	4.178***	4.223***	4.070***
	(1.242)	(1.254)	(1.258)	(1.270)
Sales_Exportsh			-1.118	
			(1.258)	
Exportshare			-0.370***	
			(0.0618)	
Exporting		0.119***		0.119***
		(0.0410)		(0.0410)
Techn_Pot				0.0189
				(0.0423)
Firm Controls	Yes	Yes	Yes	Yes
Clustering	Firm	Firm	Sector×Year	Sector
No of Clusters	109	109	109	109
Observations	12,829	12,820	12,820	12,820
R-squared	0.384	0.383	0.385	0.383

Notes: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent. Clustered standard errors in brackets. All regressions include year fixed effects, industry fixed effects and firm controls of size, ownership, age, location and market competition. $Sales_{j,t,t+4}$ is defined as the average annualized change in ownership rates between t and $t+4$.

Table 6: Robustness Check: Positive New Product Sales

Dependent Variable: Dummy for Positive New Product Sales

	(1)	(2)	(3)	(4)
$Sales^{actual}_{j,t,t+4}$	0.743***	0.761***	0.770***	0.723***
	(0.236)	(0.244)	(0.224)	(0.249)
Sales_Exportsh			-0.208	
			(0.223)	
Exportshare			-0.0714***	
			(0.0128)	
Exporting		0.107***		0.107***
		(0.00900)		(0.00897)
Techn_Pot				0.00537
				(0.00428)
Firm Controls	Yes	Yes	Yes	Yes
Clustering	Firm	Firm	Sector×Year	Sector
No of Clusters	96	96	96	96
Observations	18,313	18,254	18,254	18,254

Notes: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent. Clustered standard errors in brackets. All regressions include year fixed effects, industry fixed effects and firm controls of size, ownership, age, location and market competition. $Sales_{j,t,t+4}$ is defined as the average annualized change in ownership rates between t and $t+4$.

Regressions on the Sector Level

Next, let us tackle the potential problem of high correlation between the error terms on the firm level. Since our innovation measure comes from the firm level dataset but the market size effect is identified on the more aggregate sector level, there exists the risk of artificially low standard errors (especially if we cluster on the firm level). As mentioned above, clustering on the aggregate sector level would be the conservative approach to estimate the standard errors. However, being endowed with an insufficient number of clusters this approach is not very reliable to our purposes. Alternatively, we collapse all firm-level observations on the sector level and re-run our baseline regressions on the sector level. This should alleviate the potential problem of high correlation among error terms on the firm level.

Table 7 displays our estimates of market size on the sector level for all three dependent variables of innovation performance. All specifications include all control variables of size, age, region, market competition, ownership structures, export behaviour and technology potential.[48] To control for the different number of firms we observe in each sector, all regressions are weighted by the number of observations.

Table 7: Sector Level Regression

	(1)	(2)	(3)
Dep. Variable	Log Labour Productivity	Log Investment	Dummy Newproduct Output
$Sales^{actual}_{j,t,t+4}$	4.863***	8.851***	0.507***
	(0.0617)	(0.165)	(0.0206)
Firm Controls	Yes	Yes	Yes
Std Errors	Robust	Robust	Robust
Observations	21,853	21,853	18,323
R-squared	0.944	0.942	0.947

Notes: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent. Robust standard errors in brackets. All regressions include year fixed effects, industry fixed effects and the industry mean of all firm controls of size, ownership, age, location, market competition and export behaviour. Further they control for technology potential on the industry level. Sector observations are weighted by the number of firms (with non-missing values of each variable in each year). $Sales_{j,t,t+4}$ is defined as the average annualized change in ownership rates between t and $t+4$.

[48] We take the (unweighted) mean of all variables at the sector level (including the mean of dummies such as ownership and left hand side variables.

3.3 Accounting for Endogeneity of Actual Market Size

So far, the identification of the effect of market size on innovation activities relies crucially on the the assumption that market size measured on the sector level from household surveys is truly exogenous to the error term in the regression models (4) to (6), i.e. innovation activities are not affecting market size. As we measure innovation activities on the firm level from household survey data and concentration in all durable good sectors is quite low, we think that the exogeneity assumption might not be completely unjustified.[49]

However, as the endogeneity of market size is a major concern in any investigation of market size on innovation, we address this issue here employing a instrumental variables approach. The concern is that innovation activities of firms change the price and quality of durable goods which induces consumers to acquire durable goods for a given distribution of incomes. To account for the potential endogeneity of market size measure, we need instruments that drive the market size of each durable good but do not affect the innovation activities of firms directly, i.e. a valid exclusion restriction. We use two instruments and run the following first stage (Equation 7) and second stage (Equation 8):

$$Sales^{actual}_{j,t,t+4} = \beta_1 Sales^{potential}_{j,t,t+4} + \beta_2 Sales^{actual}_{j,t-4,t} + \mathbf{X}'_{i,j,t}\beta + \lambda_t + \eta_j + \mu_{i,j,t} \tag{7}$$

$$Inno_{i,j,t} = \alpha_1 Sales^{actual}_{j,t,t+4} + \mathbf{X}'_{i,j,t}\alpha + \eta_j + \lambda_t + \epsilon_{i,j,t} \tag{8}$$

As our first instrument, we use our measure for potential future market size, $Sales^{potential}_{j,t,t+4}$, as a predictor of actual future market size. As explained in Section 2.4, we follow Acemoglu and Linn (2004) in the construction of this potential market size measure, which is only driven by future changes in the income distribution and thus, by construction, orthogonal to price or quality changes which would induce changes in ownership patterns and create a larger market size.[50]

[49] Cf. Table 16 in the data appendix for market concentration in different sectors.

[50] Here the exclusion restriction is $Cov\left(Sales^{potential}_{j,t,t+4},\ \epsilon_{i,j,t}\right) = 0$. Of course, it could be argued, that there is still reverse causality in the sense that innovation activity in the aggregate influences output growth which ultimately trickles down to household income growth. We think, however, that this is not a major concern as the market size effect on innovation is identified on two dimensions, sectors and time, whereas it cannot be credibly claimed that this is true also in the reverse direction, i.e. innovation activities in a certain sector particularly affecting the income of those households which are most prone to acquire the goods supplied by this sector. For instance this would mean that innovation activity in the bicycle sector affects mainly the household most prone to buy new bicycles, i.e. the poorer households, whereas the innovation activity of the automobile sector would influence the income of richer households

Of course, this potential market size measure might be calculated with a error as not only income alone is affecting the evolution of durable good sales but also other factors such as urbanisation, the housing market and public good provision.[51] For this reason, we include the lag of actual observed market size, $Sales^{actual}_{j,t-4,t}$, as second excluded instrument in our first stage and a natural predictor of actual future market size. As outlined in Section 2.1, actual market size follows a particular pattern as durable goods ownership goes from take-off to saturation in the population with each phase persisting for several years. We use this natural persistence in the data to predict future market size. With regard to the exclusion restriction, past actual market size should not affect innovation activities of firms directly other than through influencing firms' expectations about the development of future market size.[52]

We first present the first stage results of instrumental variables regression of the logarithm of labour productivity on actual market size in Table 8. Column (1) shows the effect of both instruments on actual market size clustering the standard errors on the firm level. Both instruments are significantly correlated and show an effect of considerable magnitude on the endogenous variable. A one percentage point increase in predicted potential sales increases future actual sales by 0.7 percentage points whereas a one percentage point increase in lagged actual sales increases future actual sales by 0.6 percentage points. To account for the possible correlation of regression errors within sectors, we also report first stage results with sector×year clustering (in Column (2) and with observations collapsed at the sector level (Column (3)). The coefficient of lagged actual market size remains significant throughout, whereas potential market size, $Sales^{potential}_{j,t,t+4}$, loses significance (although only narrowly so in Column (2)). Statistical significance is also reflected in the F-statistic, calculated for the ex-

who are most prone to buy new cars. Thus, it is more credible that sectoral innovation activities are not influencing only certain part of the income distribution but more the income distribution in general which would not confound our identification to a large extent.

[51] In fact, empirical evidence provided by Beerli (2010) has shown that there is a considerable trend in the probability of durable goods ownership across the income distribution, i.e. the likelihood of being an owner increases considerably for all income groups. The analysis shows that the decline in durable good prices might be an important force, but also other factors such as urbanisation, public goods provision such as electricity and running water may play an important role to explain these trends.

[52] The exclusion restriction for past sales is $Cov\left(Sales^{actual}_{j,t-4,t},\ \epsilon_{i,j,t}\right) = 0$. One possible channel, when this might be violated is, if a larger market size in the past relaxes firms' liquidity constraint and allowing them to invest more in R&D compared to firms with smaller market size in the past. As we exploit variation of innovation activities on the firm level and all firms in the same sector face the same market size in the past, we think that the liquidity argument is not a first order concern. We also checked the robustness of our results, however, including a firm level control of past profits (a proxy for a relaxed liquidity constraint) and found similar results.

Table 8: First Stage Results

Dependent Variable: $Sales_{j,t,t+4}$

(Dependent Variable of Second Stage: Log of labour productivity)

	Over-identified Model			Just-identified Model	
	(1)	(2)	(3)	(4)	(5)
Aggregation level	firm	firm	sector	firm	sector
$Sales^{potential}_{j,t,t+4}$	0.709	0.709	0.343		
	[0.0824]***	[0.449]	[0.416]		
$Sales^{actual}_{j,t-4,t}$	0.640	0.640	0.920	0.649	0.933
	[0.0285]***	[0.0677]***	[0.0731]***	[0.0669]***	[0.0673]***
Firm Controls	Yes	Yes	Yes	Yes	Yes
Std Errors/Clustering	Firm	Sector×Year	Robust	Sector×Year	Robust
No of Clusters	5407	95		95	
Observations	13,133	13,133	95	13,133	95
R-squared	0.965	0.965	0.992	0.965	0.992
First Stage F-Statistic	827.5	59.02	106.8	94.10	192.6

Notes: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent. Standard errors are clustered on the indicated level for firm level regressions and robust standard errors shown for sector level regressions. All regressions include year fixed effects, industry fixed effects and firm controls of size, ownership, age, location and market competition. Observations in sector level regressions are averaged over firms in each sector and weighted by the number of firms in each sector-year cell. $Sales_{j,t,t+4}$ is defined as the average annualized change in ownership rates between t and $t+4$. The dependent variable in the second stage is the logarithm of labour productivity.

cluded instruments only and reported at the bottom of the table. In all columns it is larger than the threshold value of 10 above which IV is not supposed to be subject to weak instruments critique as proposed by Staiger and Stock (1997). For comparison, we also show the first stage results of the just-identified model in Columns (4) and (5) using only our strongest instrument, lagged actual market size. We come back to the issue of weak instruments at the end of this section.

Table 9 reports our main results of the effect of market size on labour productivity. For comparison, Column (1) shows again the effect of actual market size, $Sales^{actual}_{j,t,t+4}$ using ordinary least squares.[53] As we instrument actual market size with past actual market size and our potential market size measure (over-identified case), the coefficient of market size becomes stronger showing that a one percentage point increase in future market size is associated with an increase in productivity by 6.471 percent

[53]As the 2SLS model includes the lagged actual sales we lose the first year of firm observations and adjusted the number of observations throughout so that OLS and 2SLS estimation are carried out on the same sample of firms and comparable. This is why the number of observations in this section differs to some degree from the tables in the previous section. Qualitatively, resulting coefficient estimates are similar to the previous section but can differ slightly in magnitude (particularly for investment).

(Column (2) and (3) with firm and sector$\times$year clustering) and by almost 6 percent if observations are collapsed on the sector level. We suspect that the potential endogeneity of actual market size leads to a downward bias of OLS estimates which would explain the difference in the magnitude to 2SLS coefficients.[54]

It is well known that the 2SLS estimator is most biased when instruments are weak or when there are many over-identifying restrictions. As proposed by Angrist and Pischke (2009) we compare our results to two alternative estimation procedures to check robustness. One strategy is to pick the strongest instrument and show just-identified estimates as just-identified IV is median-unbiased. In our case, IV estimates using only lagged actual market size as an instrument for actual market size are of a similar magnitude (Column (5) for firm level and Column (6) for sector level estimates).

Furthermore, we compare our IV estimates with those produced by the limited information maximum likelihood (LIML) estimator in the over-identified case.[55] Column (7) corresponds to column (3) except for the fact that parameters are estimated with LIML rather than 2SLS. With a coefficient of 6.472 (standard error 1.312) LIML yields almost the same point estimate as 2SLS. Similar results are found comparing the LIML estimates on the sector level (Column (8)) with its 2SLS counterpart in column (4). In both cases, LIML estimates are very similar to 2SLS estimates (and a lit less precise as presumed). The comparison between 2SLS and LIML suggests that finite-sample bias is not strong concern.

In the data appendix, we also report the results of our other outcome variables which parallel the findings with respect to labour productivity. Table 18 shows second stage results for the market size effect on the probability of positive new product output. Column (1) again shows the OLS estimates from the previous section for comparison. In this case, a one percentage point increase in durable good sales in the next 5 years leads to a 0.75 percentage point increase in the probability of a firm having positive new product sales. Similar to observations with labour productivity, the magnitude of the effect increases as we move from OLS to 2SLS in the over-identified case to 1.07 percentage points (Column (2) with clustering on the firm level and Column (3) with clustering on sector$\times$year level). The effect is again a bit smaller collapsing observations on the sector level. LIML shows quite

[54]In fact, technical considerations show that the endogeneity bias of OLS estimates could go in both directions, downward or upward. Intuitively, the effect of innovation on market size can either attenuate or reinforce the effect of market size on innovation depending on the relative magnitude of each effect (Hayashi, 2000). Another possible source of the downward bias of the OLS estimate is measurement error in actual market size.

[55]The issue of weak instruments becomes more relevant with the number of weak instruments. In finite samples, LIML is less biased than 2SLS but has a higher variance.

similar results to over-identified 2SLS and results are slightly higher in the just identified case.[56] In Table 20 we present results for the logarithm of investment. Here the market size effect increases from 2.75 percent in the OLS case to 4.4 percent in the over-identified 2SLS case (Columns (2) and (3)) or to 5.38 percent in the just-identified 2SLS estimates (Column (5)). If we run regression on the sector level instead, the market size effect increases to around 8.4 percent (Columns (4) and (8) in the over-identified case) or 8.5 percent (Column (6) in the just-identified case).[57]

[56] Furthermore, an inspection of first stage results parallels largely the results found with respect to labour productivity. F statistics are generally high and above the critical threshold of 10. Cf. Table 19 in the data appendix.

[57] Cf. Table 21 in the data appendix for the first stage estimates when the logarithm of investment is the dependent variable in the second stage.

Table 9: Second Stage Results

Dependent Variable: Log of Labour Productivity

	OLS	Over-identified 2SLS			Just-identified 2SLS		LIML	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Aggregation level	firm	firm	firm	sector	firm	sector	firm	sector
$Sales_{t,t+4}$	4.986	6.471	6.471	5.940	6.719	5.996	6.472	5.955
	[1.076]***	[1.420]***	[1.312]***	[0.979]***	[1.279]***	[0.981]***	[1.312]***	[0.982]***
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Std Errors/Clustering	Sector×Year	Firm	Sector×Year	Robust	Sector×Year	Robust	Sector×Year	Robust
No of Clusters	95	5407	95		95		95	
Observations	13,133	13,133	13,133	95	13,133	95	13,133	95
R-squared	0.200	0.199	0.199	0.945	0.199	0.945	0.199	0.945

Notes: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent. Standard errors are clustered on the indicated level for firm level regressions and robust standard errors shown for sector level regressions. All regressions include year fixed effects, industry fixed effects and firm controls of size, ownership, age, location and market competition. Observations in sector level regressions are averaged over firms in each sector and weighted by the number of firms in each sector-year cell. $Sales_{j,t,t+4}$ is defined as the average annualized change in ownership rates between t and $t+4$. In the over-identified model and LIML two instruments, $Sales^{actual}_{j,t-4,t}$ and $Sales^{potential}_{j,t,t+4}$, are used to instrument for $Sales_{j,t,t+4}$ whereas the just-identified model uses only $Sales^{actual}_{j,t-4,t}$.

Discussion and Future Research

Although the high values of the F-statistic indicate that both instruments are jointly significant and have some power in the prediction of actual future market size, the coefficient of potential market size, our preferred instrument, loses significance if we cluster standard errors in a more demanding way or estimate the effect collapsing observations on the sector level (even though the point estimate of potential market size remains positive throughout). Potential market size, as constructed and explained in Section 2.4, is our preferred instrument for actual market size, as its time dynamic is only driven by shifts in the income distribution. Thus, by construction the time dynamic in potential market size is orthogonal to price and quality changes of durable goods induced by firms' innovation activities which would induce usage intensities to change for a given income group.[58]

Now, Beerli (2010) shows in a descriptive analysis of the dynamic of durable goods ownership that income is an important predictor of durable goods ownership but that there are also other important influence factors, such as public infrastructure, urbanisation and special features in the dynamic of the Chinese housing market, which drive the dynamic of durable goods ownership apart from income and durable goods price changes. We suspect that the potential market size measure, as calculated here, might neglect part of the dynamic stemming from these other influence factors which are also orthogonal to firms' innovation activities in our sense. Neglecting part of this dynamic could result in the weaker performance of the potential market size instrument in the first stage. Furthermore, it is also likely that part of these other influence factors are captured in the other instrument, the lagged actual market size, which would absorb a considerable share of the variation leaving the coefficient of potential market size insignificant.

Another drawback of the identification strategy employed here is that we have only 98 data points to identify the effect of market size (as measured on the sector level) on innovation with a five-year time horizon for future market size.[59] In future work, we plan to build a richer measure of potential market size which integrates the influence factors mentioned above together with the dynamics in the income distribution. Additionally, we currently work on an instrument which exploits the geographical dimension in our firm and household data set. Particularly, we construct a province

[58]Recall, potential market size is calculated as $Sales^{potential}_{j,t,t+4} = \sum_g \bar{u}_{j,g} \left(i_{g,t+4} - i_{g,t}\right)$. Thus, the aggregate dynamic of $Sales^{potential}_{j,t,t+4}$ solely from the two facts that, first, the population weight of income groups changes over time as all households become richer and that, second, preferences differ across income groups (i.e. they are non-homothetic) but remain stable over time. Thus, price and quality changes induced by firms' innovation activities cannot influence the $\bar{u}_{j,g}$ of a given income group by construction.

[59]With a five year time horizon we lose two sectors and three years and end up with a panel of 14 different durable good sectors and 7 years.

level market size measure with variation on a third dimension apart form sector and time. Beerli (2010) showed that there are considerable geographical differences in durable good ownership as rich coastal provinces approach saturation in most lower to middle ranking durable goods in recent years whereas car acquisitions are booming. In contrast, inland provinces lag behind showing higher durable good acquisitions in low and middle ranking durable goods and still low acquisitions in cars. We expect the exploitation of the geographical dimension of market size to be an interesting next step in our analysis of the effect of market size on innovation activities of Chinese firms.

4 Conclusion

In the last three decades, China witnessed a formidable growth miracle creating a surging middle class of new consumers with discretionary income to spend on consumer goods. In this paper, we investigate the response of innovation activities across different Chinese manufacturing firms to changes in market size of household durable goods driven by these large changes in the Chinese income distribution. In so doing, we link market size information inferred from changes in ownership patterns from a long panel of household surveys (i.e. the China Health and Nutrition Survey (CHNS)) to the universe of Chinese manufacturing firms taken from the Annual Survey of Industrial Production (ASIP).

In the first part of the paper, we document how the Engel properties of consumer preferences for durable goods (i.e. richer income groups exhibiting an over-proportional taste for relative luxurious goods) and changes in the income distribution are driving the dynamics in the acquisitions across different durable goods. Over the last two decades, the sales in durable goods have shifted away from relative necessities (e.g. bicycles and electric fans) to more higher ranking goods (e.g. air conditions and cars).

In the second part of the paper, we establish a (causal) relation between changes in durable goods sales and innovation activities of Chinese manufacturing firms making use of an instrumental variables strategy. As the observed actual market size is likely to be endogenous, with qualitatively improved and cheaper products having larger markets, we follow Acemoglu and Linn (2004) and construct a measure of potential market size as an instrument for actual market size. Exploiting information on average ownership of durable goods across income groups in the household data, we fix ownership of income groups to a specific base-year and use only the changes in the population share of the income groups across time to infer the dynamic in potential market size of different durable goods from 1998 to 2005. Constructed like this, price and quality changes resulting from innovation activities of Chinese firms do not affect potential market size. We complement this measure for potential market size with lagged actual market size which should also be exogenous to contemporaneous innovation activities and captures the natural persistence in the dynamics of sales.

Looking at different measures of innovation activities on the firm level, we find that a one percentage point increase in market size over the next five years raises the probability of successful product innovation by about 1.1 percentage points, labour productivity by 6.5% and R&D inputs (as measured by investment) by 4.4%. These results were found to be robust to including a rich set of firm-level determinants of R&D and the sector market concentration. Additionally, we show that controlling for export behaviour of firms does not affect the results addressing the concern that the non-domestic market might confound the effect of domestic market size on innovation activities. Furthermore, we include a measure of worldwide technology potential reported by Swiss firms in our regression framework to demonstrate that results are also robust to supply side drivers of R&D.

To the best of our knowledge, this paper is the first empirical attempt to identify a causal relation between differential technical change and changes in market size driven by the Engel curve property of consumer preferences across a broad scale of manufacturing industries. In so doing, this paper generalises the evidence found by Acemoglu and Linn (2004) on the pharmaceutical industry to a wider set of different manufacturing industries. Furthermore, the evidence shown here provides a unifying picture for the literatures on both directed technical change and demand driven technical change, both of which stress the importance of market size and profit incentives in innovation activities. If the market size hypothesis is true as suggested by our findings, (Chinese) firms operating in markets of relative luxurious consumer goods will see a rise in their productivity while firms producing necessities will witness a decline in their productivity along the course of economic growth.
In future research we plan to exploit the geographical information available in our firm and household data set to construct a measure of province level market size. Apart from providing a market size measure with a third dimension of variation (in addition to sectors and years), we expect this to yield interesting new insights as there are considerable differences in economic development and household income across Chinese coastal and inland provinces.

References

Acemoglu, D. (1998), "Why Do New Technologies Complement Skills? Directed Technical Change and Wage Inequality", *The Quarterly Journal of Economics*, 113, 1055-1089.

Acemoglu, D. (2002), "Directed Technological Change", *Review of Economic Studies*, 64, 781-809.

Acemoglu, D., Aghion, Ph., Griffith, R. and Zilibotti, F. (2010), "Vertical Integration and Technology: Theory and Evidence", *Journal of European Economic Association*, 8 (5), 989-1033.

Acemoglu, D. and Linn, J. (2004), "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry", *The Quarterly Journal of Economics*, 119, 1049-1090.

Acemoglu, D. and Zilibotti, F. (2001), "Productivity Differences", *Quarterly Journal of Economics*, 66, 563-606.

Ackerberg, D.A., Caves, K., and Frazer, G. (2006), "Structural Identification of Production Functions", Mimeo, UCLA.

Aghion, P. and Howitt, P. (1992), "A Model of Growth through Creative Destruction", *Econometrica*, 60, 323-351.

Aghion, P. and Howitt, P. (1992), *Endogenous Growth Theory*, MA: MIT Press.

Aghion, P., Bloom, N., Blundell, R., Griffith, R. and Howitt, P. (2005), "Competition and Innovation: An Inverted-U Relationship", *The Quarterly Journal of Economics*, May 2005, 701-728.

Angrist, J. and Pischke, J. (2009), *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press.

Arnold, J. M. and Katrin Hussinger (2005), "Export Behaviour and Firm Productivity in German Manufacturing: A Firm-Level Analysis", *Review of World Economics*, 141, 219-243.

Bayus, B. and Gupta, S. (1992), "An Empirical Analysis of Consumer Durable Replacement Intentions", *International Journal of Research in Marketing*, 9 (3), 257-267.

Beerli, A. (2010), "The Evolution of Durable Goods Demand During China's Transition - An Empirical Analysis of Household Survey Data from 1989 to 2006", IPCDP Working Paper, 201.

Ben-David, D. (1993), "Equalizing Exchange: Trade Liberalization and Income Convergence", *Quarterly Journal of Economics*, 108, 653-679.

Boppart, T. and Weiss, F. J. (2012), "Structural Change, Market Size and Sector Specific Endogenous Growth", Mimeo, University of Zurich.

Boppart, T. (2011), " Structural Change, Relative Price Effects and the Kaldor Facts in a Growth Model with Non-Gorman Preferences", IPCDP Working Paper 203.

Brandt, L., Rawski, T. and Sutton, J. (2008), "China's Industrial Development", ch.15, 569-632. In *China's Great Economic Transformation*, edited by Brandt, L. and Rawski, T., Cambridge University Press, Cambridge.

Brandt, L., v. Biesenbroeck, J. and Zhang, Y. (2011), "Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing", *Journal of Development Economics*, 97 (2), 339-351.

Crépon, B., Duguet, E. and Mairesse, J. (1998), "Research, Innovation and Productivity: An Econometric Analysis at the Firm Level", *Economics of Innovation and New Technology*, 7, 115-158.

CHNS (2012). 2009 Household Survey. Accessed 11.5.2012, *www.cpc.unc.edu/projects/china/data/questionnaires/C09hh_Fin20090710.pdf* (accessed: 31.5.2012)

Cotterill, R. W. (1986), "Market Power and the Retail Food Industry: Evidence from Vermont", *The Review of Economics and Statistics*, 68, 379-386.

Engel, E. (1857), "Die Productions- und Consumptionsverhaeltnisse des Koenigsreichs Sachsen", *Zeitschrift des Statistischen Bureaus des Königlich Sächsischen Ministeriums des Inneren*, No. 8 and 9.

Farrell, J. and Shapiro, C. (1990), "Asset Ownership and Market Structure in Oligopoly", *The RAND Journal of Economics*, 21, 275-292.

Farrell, J. and Shapiro, C. (1990), "Horizontal Mergers: An Equilibrium Analysis", *The American Economic Review*, 80, 107-126.

Feldman, S., McClain, D. and Palmer, K. (1987), "Sources of Structural Change in the United States, 1963-1978: An Input-Output Perspective", *The Review of Economic Studies and Statistics*, **69**(3), 503-510.

Foellmi, R. and Zweimüller, J. (2006), "Income Distribution and Demand-Induced Innovations", *Review of Economic Studies*, 73, 941- 960.

Foellmi, R. and Zweimüller, J. (2008), "Structural Change, Engel's Consumption Cycles and Kaldor's Facts of Economic Growth", *Journal of Monetary Economics*, 55, 1317-1328.

Grossman, G. and Helpman, E. (1991), "Innovation and Growth in the Global Economy", MIT Press, MA.

Habakkuk, H.J. (1962), *American and British Technology in the Nineteenth Century: Search for labour Saving Inventions*, Cambridge: Cambridge University Press.

Herrendorf, B., Rogerson R. and Valentinyi, A. (2009), "Two Perspectives on Preferences and Structural Transformation", NBER Working Paper, 15416.

Hu, A. and Jefferson, G., (2008) "A great wall of patents: what is behind China's recent patent explosion?", *Journal of Development Economics*, 90, 57-68.

Jefferson, G., Hu, A., Guan, X. and Yu, X. (2003), "Ownership, performance, and innovation in China's large- and medium-size industrial enterprise sector", *China Economic Review*, 14, 89-113.

Jefferson, G., Huamao, B., Xiaojing, G. and Xiaoyun, Y. (2006), "R&D Performance in Chinese Industry", *Economics of Innovation and New Technology*, 15, 345 -366.

Klepper, St. (1996). "Entry, Exit, Growth, and Innovation over the Product Life Cycle." *The American Economic Review*, 86 (3), 962 - 983.

Konjunkturforschungsstelle (KOF) ETH Zürich (2012). *http://www.kof.ethz.ch/en/surveys/structural-surveys/innovation-survey* (accessed: 31.5.2012)

Kongsamut, P., Rebelo, S. and Xie, D. (2001), "Beyond Balanced Growth", *Review of Economic Studies*, 68, 869-882.

Matsuyama, K. (1992), "Agricultural Productivity, Comparative Advantage, and Economic Growth", *Journal of Economic Theory*, 58, 317-334.

Mairesse, J. and Mohnen, P. (2010), "Using Innovations Surveys for Econometric Analysis", NBER Working Paper 15857.

National Bureau of Statistics (NBS) (2008), Guo min jing ji hang ye fen lei zhu shi. Zhong guo tong ji chu ban she, Beijing.

Olley, G. S. and Pakes, A. (1996), "The dynamics of productivity in the telecommunications equipment industry", *Econometrica*, 64, 1263-1297.

Staiger, D. and Stock, J. (1997), "Instrumental Variables Regression with Weak Instruments", *Econometrica*, 65(3), 557-586.

Schmookler, J. (1966), "Invention and Economic Growth", Harvard University Press, Cambridge, MA.

Song, Z., Storesletten, K. and Zilibotti, F. (2011)."Growing Like China", *The American Economic Review*, 101, 202-241.

World Bank (2009). Gross national income per capita 2009, Atlas method and PPP. Available at: *http://www.damtp.cam.ac.uk/user/na/FoCM11/GNIPC.pdf* (accessed 31.5.2012)

World Bank (2012a). How We Classify Countries.
http://data.worldbank.org/about/country-classifications (accessed 31.5.2012)

World Bank (2012b). World dataBank. World Development Indicators (WDI) & Global Development Finance (GDF)
http://data.worldbank.org/country/china (accessed: 31.5.2012)

Data Appendix

A.1 Construction of Potential Market Size Measure

Definition of Income Groups

Household income and household income per capita is provided by the CHNS in longitudinal data-files including the latest wave 2009.[60] Household disposable income in the CHNS is conceptualised as the sum of all sources of market and non-market incomes or revenues minus expenses on the household or individual level. We use houehold income delfated to constant 2009 Yuan, using the price deflator also provided by the CHNS (2012a,b) which is based on a standard NBS consumer basket allowing for price differences between urban and rural areas.

We split the income distribution into $g = 1, ..., G$ groups setting fixed income thresholds in constant 2009 Yuan and calculate the population share $i_{g,t}$ of each income group g for each survey year t. We use different kinds of income group definitions using stable income thresholds to check robustness

In our baseline, we take inspiration from the World Bank's (WB, 2012) classification of countries[61] and divide households into four ($G = 4$) income groups: low income, lower middle income, upper middle income and high income. The World Bank's thresholds in constant 2010 dollars and were converted into constant 2009 yuan. The threshold for the high income group was substantially lowered (from originally $12'276) in order to get an accurate measure of its group size in the CHNS wave 1997. All dollar figures where converted into constant 2009 Yuan using the exchange rate and PPP adjustment factors.[62]

Usage Profiles and Base-Year

Further we calculate the average usage intensities, $u_{j,g}$, i.e. the number of items of durable good j per household, for each income group and all durable goods. In principle, we could calculate these average usage intensities of each income group durable good pair for all survey years available in the CHNS, i.e. we have $u_{j,g,t}$ for all survey years.

[60]See Beerli (2010) for a more detailed description and a comparison offical household survey by the NBS.

[61]The World Bank classifies economies according to their 2009 GNI per capita, calculated using the World Bank Atlas method. The following thresholds are set: low income, $1'005 or less; lower middle income, $1'005 - $3'975; upper middle income, $3'975 - $7'675; and high income, $7'675 or more.

[62]Dollar values are converted to constant 2009 using the China Version 2 exchange rate (6.83) and PPP adjustment factor (3.16) from the Penn World Tables, i.e. $threshold \times \frac{XRAT}{PPP}$. This yields the following thresholds in constant 2009 Yuan: low income (2149 Yuan), lower middle income (2150 - 8515 Yuan), upper middle income (8516 - 16'500 Yuan), high income (16'500 or more).

In order to abstain from ownership changes that are not driven by income, we need to decide on a base-year from which we take the usage profiles, $u_{j,g,t=2009} \equiv \bar{u}_{j,g}$, for all income groups to calculate the aggregate, potential ownership rates in other years. Given a base-year, the aggregate stock a durable good j and the annualised sales, i.e. the flow, are given by

$$Stock_{j,t}^{potential} = \sum_{g} \bar{u}_{j,g} i_{j,t} \tag{9}$$

$$Sales_{j,t,t+k}^{potential} = \left[Stock_{j,t+k}^{potential} - Stock_{j,t}^{potential}\right] \frac{1}{k} = \left[\sum_{g} \bar{u}_{j,g} \left(i_{g,t+k} - i_{g,t}\right)\right] \frac{1}{k} \tag{10}$$

The choice of a base-year for ownership profiles implies different assumptions about entrepreneurs expectations, on the one hand, and accuracy considerations on the other hand. Taking ownership profiles from a survey year at the beginning of our panel of manufacturing firms, e.g 1997, we assume that entrepreneurs base their R&D decisions on expectations about future market evolution based on information about ownership profiles conditional on durable good prices and qualities from 1997, i.e. one year before our firm panel starts. As Beerli (2010) shows in his analysis of durable good ownership between 1989 and 2006, depending on the durable good, ownership rates were generally increasing across the income distribution mainly explained by a substantial fall in durable goods prices but also by improvements in public service provision and other factors. Additionally, ownership rates increased unevenly across the income distribution with poor households gaining much more from price changes compared to richer income groups. This implies that the aggregate, potential ownership stocks calculated based a base-year 1997 will turn out to be lower then it actually is. With respect to accuracy, picking 1997 as a base-year involves the problem that there are relatively few rich households (i.e. less than 1%) which makes the information about their ownership profiles relatively inaccurate.[63] Taking the latest survey year available, i.e. 2009, on the other hand, implies assuming that entrepreneurs form their expectation about the future development of durable good sales based on ownership information conditional on good quality and durable prices that are generally much lower at the end of the firm survey. Thus, we aggregate, potential durable good stocks will be overestimated for earlier years. Yet, since there are many more rich households in 2009 than in earlier years, their ownership profile are relatively more accurate. Thus, independently from the choice of the base-year, potential stocks will be either over- or underestimated throughout the panel which means that potential sales, the difference between two years, will generally be lower than actual sales.[64]

[63]Another problem is that some durable good become available only in later survey years, e.g. cell phones from 2004.

[64]This is in line with the findings of Beerli (2010) who finds that the share of changes in aggregate ownership

A.2 Tables

Table 10: Four Year Averages of $Sales_{t,t+1}$

Durable Good	**1989-1993**	**1994-1998**	**1999-2003**	**2004-2008**
air condition	0.004	0.013	0.027	0.043
bicycle	0.007	-0.038	-0.064	-0.035
camera	0.008	0.006	0.003	0.003
car	0.002	0.002	0.002	0.005
cellphone				0.158
colour TV	0.036	0.050	0.043	0.026
computer		0.009	0.012	0.034
cycle	0.016	-0.032	-0.058	-0.039
dvd			0.036	0.000
electric fan	0.097	0.038	0.020	0.013
refrigerator	0.022	0.027	0.016	0.044
homevideo appliances	0.036	0.084	0.090	0.027
kitchen appliances	0.066	0.073	0.061	0.082
microwave	0.002	0.008	0.022	0.021
motorcycle	0.008	0.025	0.019	0.014
presscooker	0.029	0.028	0.005	0.006
radio	0.030	-0.009	-0.029	-0.057
ricecooker	0.035	0.037	0.033	0.055
satellite dish				0.015
sewing machine	-0.006	-0.008	-0.018	-0.017
telephone		0.073	0.054	-0.018
tricycle	0.009	0.006	0.005	-0.003
washing machine	0.014	0.022	0.013	0.029

Notes: The table shows $Sales_{t,t+1}$ averaged over four consecutive years for each durable good.

explained by income can differ substantially between different durable goods, being only 31% for colour TVs.

Table 11: Durable Good Stocks in CHNS Panel 1989 - 2009

Durable Good	**1989**	**1991**	**1993**	**1997**	**2000**	**2004**	**2006**	**2009**
air condition	0.004	0.003	0.014	0.058	0.102	0.221	0.288	0.437
bicycle	1.410	1.427	1.487	1.321	1.223	0.936	0.839	0.760
camera	0.058	0.069	0.091	0.108	0.136	0.144	0.147	0.158
car	0.007	0.010	0.015	0.024	0.030	0.039	0.046	0.066
cellphone						0.657	0.918	1.446
colour TV	0.191	0.250	0.326	0.503	0.681	0.838	0.903	0.970
computer				0.027	0.054	0.103	0.140	0.271
cycle	1.443	1.481	1.559	1.419	1.336	1.072	0.985	0.879
dvd					0.256	0.402	0.414	0.403
electric fan	0.891	1.135	1.321	1.539	1.582	1.668	1.687	1.734
refrigerator	0.126	0.171	0.210	0.319	0.400	0.454	0.506	0.672
homevideo appliances	0.191	0.250	0.326	0.503	0.934	1.240	1.317	1.373
kitchen appliances		0.477	0.601	0.903	1.107	1.344	1.543	1.752
microwave		0.002	0.005	0.019	0.060	0.154	0.193	0.258
motorcycle	0.022	0.024	0.041	0.135	0.213	0.284	0.318	0.352
presscooker		0.278	0.334	0.455	0.532	0.533	0.570	0.565
radio	0.384	0.534	0.541	0.521	0.479	0.350	0.236	.
ricecooker		0.200	0.262	0.430	0.522	0.657	0.780	0.930
satellite dish							0.065	0.110
sewing machine	0.548	0.540	0.525	0.488	0.470	0.388	0.345	0.302
telephone				0.290	0.509	0.704	0.687	0.614
tricycle	0.033	0.054	0.073	0.102	0.115	0.136	0.146	0.119
washing machine	0.344	0.374	0.392	0.488	0.546	0.592	0.650	0.738

Table 12: Usage Profiles, $u_{j,g}$, of Income Groups According to WB (2009) Classification, Base-year 2009

	Usage intesity in income group (Increase in usage intensity from lower group)				Income Group
Durable Good	**Low**	**Low Middle**	**High Middle**	**High**	**with Largest Increase**
air conditioner	0.198	0.275	0.477	0.851	
		(0.077)	(0.202)	(0.375)	High
camera	0.049	0.078	0.178	0.361	
		(0.029)	(0.101)	(0.183)	High
car	0.032	0.048	0.063	0.126	
		(0.016)	(0.015)	(0.063)	High
cellphone	1.037	1.361	1.571	1.757	
		(0.324)	(0.210)	(0.187)	Low Middle
computer	0.108	0.177	0.306	0.525	
		(0.068)	(0.130)	(0.218)	High
cycle	0.645	0.862	0.986	0.951	
		(0.218)	(0.123)	(-0.0348)	Low Middle
electric fan	1.430	1.782	1.792	1.768	
		(0.352)	(0.011)	(-0.0239)	Low Middle
homevideo appliances	1.157	1.331	1.432	1.536	
		(0.173)	(0.101)	(0.104)	Low Middle
kitchen appliances	1.240	1.547	1.909	2.279	
		(0.308)	(0.362)	(0.370)	High
motorcycle	0.279	0.406	0.361	0.294	
		(0.127)	-(0.046)	(-0.066)	Low Middle
radio	0.150	0.195	0.337	0.426	
		(0.045)	(0.142)	(0.088)	High Middle
refrigerator	0.455	0.543	0.788	0.921	
		(0.089)	(0.245)	(0.133)	High Middle
satelite dish	0.107	0.128	0.090	0.110	
		(0.021)	(-0.038)	(0.021)	Low Middle
sewing machine	0.266	0.281	0.340	0.329	
		(0.016)	(0.059)	(-0.0111)	High Middle
telephone	0.392	0.521	0.705	0.832	
		(0.129)	(0.184)	(0.127)	High Middle
wash machine	0.584	0.659	0.797	0.919	
		(0.075)	(0.139)	(0.122)	High Middle

Notes: All data are from CHNS, wave 2009. Households are grouped according to household income per capita in constant in constant 2009 Yuan: low income (2150 Yuan), lower middle income (2150 - 8515 Yuan), upper middle income (8516 - 16'500 Yuan), high income (16'500 or more). The first row of each durable good shows usage intensities (the $\bar{u}_{j,g} = u_{j,g,t=2009}$s), i.e. the average number of goods per household, and the second row shows the increase in the usage intensity (in brackets) moving from the income group below into the income group of the column.

Table 13: Industry Means of Innovation Variables, 1998-2007

Industry	**New Product Innovators**	**New Product Share**	**Investment per Worker**	**Value Added per Worker**
cellphone	0.273	0.186	51.305	338.769
car	0.372	0.185	41.357	153.699
computer	0.243	0.169	30.808	264.719
telephone	0.272	0.163	16.305	150.361
refrigerator	0.265	0.159	17.657	109.375
homevideo appliances	0.238	0.159	29.356	144.319
washing machine	0.239	0.125	20.119	99.106
air condition	0.183	0.106	19.644	131.403
camera	0.162	0.081	16.211	87.265
satellite dish	0.169	0.080	6.933	71.493
motorcycle	0.197	0.072	24.791	123.564
kitchen appliances	0.163	0.069	17.444	84.395
radio	0.114	0.058	9.456	56.711
sewing machine	0.162	0.057	12.295	64.555
electric fan	0.105	0.030	8.643	69.322
cycle	0.058	0.021	5.900	55.703
total durable goods industries	0.190	0.098	19.286	112.075
total manufacturing	0.087	0.033	16.145	93.351

Notes: All data are form ASIP 1998 to 2007. Innovation variables are the mean share of new product innovators among all firms in an industry, the share of new products on total sales, investment per worker and value added per workers as labour productivity. Industries ranked according to their share of new products on total sales.

Table 14: Correspondence between CHNS Durable Good Categories and ASIP Industries

Durable Good in CHNS	**Industry Name in CIC**	**CIC pre 2003**	**CIC post 2003**
air condition	Home air conditioner manufacturers	4065	3952
bicycle	Bicycle manufacturers	3740	3741
camera	Camera and equipment manufacturing	4254	4153
car	Automobile manufactoring	3721-3725	3721
cellphone	Mobile communications and terminal equipment manufacturing	4113	4014
colour TV	Home video equipment manufacturing	4171	4071
computer	Computer machine manufacturing	4141	4041
dvd	Home video equipment manufacturing	4171	4071
electric fan	Manufacturers of household electrical appliances ventilation	4064	3953
refrigerator	Household refrigerating appliances manufacturing	4063	3951
microwave	Household kitchen appliances manufacturing	4066	3954
motorcycle	Motorcycle manufacturing	3731	3731
presscooker	Household kitchen appliances with manufacturing	4066	3954
radio	Home audio equipment manufacturing	4172	4072
ricecooker	Household kitchen appliances manufacturing	4066	3954
satellite dish	Radio and television receiving equipment manufacturing	4130	4032
sewing machine	Sewing machinery manufacturing	3674	3653
telephone	Communication terminal equipment manufacturing	4113	4013
tricycle	Bicycle manufacturing	3740	3741
washing machine	Household cleaning electrical appliances manufacturing	4061, 4062	3955

Table 15: Descriptive Statistics of Control Variables (part I)

		State Owned		Collective Owned		Private Owned		Foreign Owned	
Industry	**Nr of obs**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
air condition	2'088	0.033	0.178	0.318	0.466	0.263	0.441	0.386	0.487
camera	941	0.095	0.293	0.090	0.287	0.089	0.285	0.725	0.447
car	3'429	0.326	0.469	0.420	0.494	0.083	0.277	0.165	0.372
cellphone	1'071	0.024	0.154	0.168	0.374	0.202	0.401	0.604	0.489
computer	1'461	0.108	0.311	0.266	0.442	0.148	0.355	0.477	0.500
cycle	4'847	0.044	0.204	0.237	0.425	0.378	0.485	0.340	0.474
electric fan	1'654	0.042	0.201	0.351	0.478	0.371	0.483	0.230	0.421
refrigerator	1'211	0.069	0.254	0.382	0.486	0.323	0.468	0.224	0.417
homevideo appliances	2'607	0.076	0.265	0.222	0.416	0.192	0.394	0.508	0.500
kitchen appliances	2'411	0.007	0.086	0.207	0.405	0.498	0.500	0.288	0.453
motorcycle	2'071	0.085	0.279	0.372	0.484	0.416	0.493	0.125	0.331
radio	3'574	0.040	0.196	0.153	0.360	0.232	0.422	0.574	0.495
satellite dish	1'543	0.088	0.284	0.198	0.398	0.369	0.483	0.345	0.475
sewing machine	2'250	0.058	0.233	0.290	0.454	0.427	0.495	0.225	0.418
telephone	1'737	0.132	0.339	0.233	0.423	0.172	0.378	0.459	0.498
washing machine	1'429	0.036	0.186	0.288	0.453	0.364	0.481	0.313	0.464
total durable good industries	34'324	0.085	0.278	0.266	0.442	0.289	0.453	0.359	0.480
total manufacturing	1'925'846	0.099	0.298	0.319	0.466	0.372	0.483	0.208	0.406

Table 16: Descriptive Statistics of Control Variables (part II)

	Firm Size		**Age**		**Exporting (yes/no)**		**Exportshare**		**Herfindahl Index**	
Industry	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**	**Mean**	**SD**
air condition	597.649	2’110.868	10.913	72.439	0.369	0.483	0.147	0.292	957.816	302.741
camera	654.965	1’188.742	8.670	8.550	0.736	0.441	0.582	0.436	852.814	190.156
car	1’586.481	5’472.603	22.156	105.140	0.200	0.400	0.018	0.092	526.271	111.251
cellphone	1’013.773	4’061.420	6.122	6.006	0.513	0.500	0.280	0.382	1’147.173	242.208
computer	1’330.065	7’312.768	7.401	8.592	0.385	0.487	0.261	0.412	646.835	308.172
cycle	253.932	495.052	12.083	67.281	0.496	0.500	0.303	0.401	144.405	40.984
electric fan	425.099	920.747	13.097	82.951	0.467	0.499	0.333	0.425	651.723	640.660
refrigerator	696.837	1’980.013	11.885	77.546	0.374	0.484	0.133	0.272	1’560.881	620.306
homevideo appliances	859.743	2’330.469	8.254	38.557	0.591	0.492	0.405	0.434	430.007	221.618
kitchen appliances	399.817	1’655.464	6.167	5.718	0.511	0.500	0.349	0.426	925.118	412.277
motorcycle	509.438	1’026.799	9.328	59.543	0.382	0.486	0.132	0.258	376.126	43.663
radio	521.761	1’031.354	8.670	45.848	0.668	0.471	0.559	0.454	298.504	125.464
satellite dish	253.658	481.770	8.342	8.000	0.469	0.499	0.323	0.416	370.253	105.399
sewing machine	227.337	334.495	10.884	42.192	0.475	0.499	0.229	0.329	354.417	96.938
telephone	536.558	1’093.505	9.996	10.877	0.444	0.497	0.294	0.413	883.710	790.562
washing machine	392.579	789.754	8.705	9.018	0.545	0.498	0.268	0.371	543.984	165.870
total durable good industries	630.231	2’712.086	10.860	57.506	0.472	0.499	0.287	0.402	563.080	459.715
total manufacturing	264.995	976.811	12.284	64.561	0.277	0.448	0.172	0.343	195.940	346.480

Table 17: Technology Potential Accross Industries

Industry	**Mean**	**SD**	**MIN**	**MAX**
air condition	2.568	0.339	2.061	2.848
camera	2.421	0.309	2.097	2.999
car	2.837	0.090	2.667	2.907
cellphone	2.349	0.108	2.276	2.507
computer	3.430	0.334	3.000	4.000
cycles	2.719	0.983	1.111	3.849
electric fan	2.526	0.352	2.061	2.848
refrigerator	2.476	0.365	2.061	2.848
homevideo appliances	2.408	0.116	2.276	2.570
kitchen appliances	2.326	0.367	2.061	2.848
motorcycle	2.550	1.050	1.111	3.849
radio	2.488	0.448	2.000	2.981
satellite dish	2.397	0.118	2.276	2.570
sewing machine	2.757	0.013	2.739	2.770
telephone	2.421	0.113	2.276	2.570
washing machine	2.509	0.363	2.061	2.848
total durable good industries	2.593	0.568	1.111	4.000

Notes: All data are form the KOF Innovation Survey (2012). Swiss firms asses the technology potential on a five point Likert-scale, i.e. the world wide availability of technological know-how in private and public hands which could be used to generate marketable, new products. Firm level assessments of technology potential was aggregated (using firm weigths) to four digit CIC and matched to the ASIP data.

Table 18: Second Stage Results

Dependent Variable: Dummy for Positive New Product Sales

	OLS	Over-identified 2SLS			Just-identified 2SLS		LIML	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Aggregation level	firm	firm	firm	sector	firm	sector	firm	sector
$Sales_{t,t+4}$	0.769	1.070	1.070	0.794	1.140	0.827	1.070	0.793
	[0.311]**	[0.507]**	[0.347]***	[0.299]***	[0.343]***	[0.293]***	[0.347]***	[0.303]***
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Std Errors/Clustering	Sector x Year	Firm	Sector x Year	Robust	Sector x Year	Robust	Sector x Year	Robust
No of Clusters	82	5343	82		82		82	
Observations	11,600	11,600	11,600	82	11,600	82	11,600	82
R-squared	0.209	0.209	0.209	0.951	0.209	0.951	0.209	0.951

Notes: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent. Standard errors are clustered on the indicated level for firm level regressions and robust standard errors shown for sector level regressions. All regressions include year fixed effects, industry fixed effects and firm controls of size, ownership, age, location and market competition. Observations in sector level regressions are averaged over firms in each sector and weighted by the number of firms in each sector-year cell. $Sales_{j,t,t+4}$ is defined as the average annualized change in ownership rates between t and $t+4$. In the over-identified model and LIML two instruments, $Sales^{actual}_{j,t-4,t}$ and $Sales^{potential}_{j,t,t+4}$, are used to instrument for $Sales_{j,t,t+4}$ whereas the just-identified model uses only $Sales^{actual}_{j,t-4,t}$.

Table 19: First Stage Results

Dependent variable: $Sales_{j,t,t+4}$

(Dependent Variable of Second Stage: Dummy for Positive New Product Sales)

	Over-identified Model			Just-identified Model	
	(1)	(2)	(3)	(4)	(5)
Aggregation level	firm	firm	sector	firm	sector
$Sales^{potential}_{j,t,t+4}$	0.677	0.677	0.430		
	[0.0944]***	[0.440]	[0.460]		
$Sales^{actual}_{j,t-4,t}$	0.650	0.650	0.862	0.660	0.877
	[0.0313]***	[0.0680]***	[0.0514]***	[0.0679]***	[0.0483]***
Firm Controls	Yes	Yes	Yes	Yes	Yes
Std Errors/Clustering	Firm	Sector×Year	Robust	Sector×Year	Robust
No of Clusters	5343	82		82	
Observations	11,600	11,600	82	11,600	82
R-squared	0.969	0.969	0.993	0.968	0.993
First Stage F-Statistic	1045	53.26	167.8	94.42	329.8

Notes: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent. Standard errors are clustered on the indicated level for firm level regressions and robust standard errors shown for sector level regressions. All regressions include year fixed effects, industry fixed effects and firm controls of size, ownership, age, location and market competition. Observations in sector level regressions are averaged over firms in each sector and weighted by the number of firms in each sector-year cell. $Sales_{j,t,t+4}$ is defined as the average annualized change in ownership rates between t and $t+4$. The dependent variable in the second stage is a dummy for positive new product sales.

Table 20: Second Stage Results

Dependent Variable: Log of Investment

	OLS	Over-identified 2SLS			Just-identified 2SLS		LIML	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Aggregation level	firm	firm	firm	sector	firm	sector	firm	sector
$Sales_{t,t+4}$	2.745	4.400	4.400	8.397	5.383	8.538	4.408	8.367
	[2.101]	[3.088]	[2.729]	[1.851]***	[2.734]**	[1.857]***	[2.734]	[1.862]***
Firm Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Std Errors/Clustering	Sector×Year	Firm	Sector×Year	Robust	Sector×Year	Robust	Sector×Year	Robust
No of Clusters	95	3819	95		95		95	
Observations	7,838	7,838	7,838	95	7,838	95	7,838	95
R-squared	0.422	0.422	0.422	0.951	0.422	0.951	0.422	0.951

Notes: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent. Standard errors are clustered on the indicated level for firm level regressions and robust standard errors shown for sector level regressions. All regressions include year fixed effects, industry fixed effects and firm controls of size, ownership, age, location and market competition.Observations in sector level regressions are averaged over firms in each sector and weighted by the number of firms in each sector-year cell. $Sales_{j,t,t+4}$ is defined as the average annualized change in ownership rates between t and $t+4$. In the over-identified model and LIML two instruments, $Sales^{actual}_{j,t-4,t}$ and $Sales^{potential}_{j,t,t+4}$, are used to instrument for $Sales_{j,t,t+4}$ whereas the just-identified model uses only $Sales^{actual}_{j,t-4,t}$.

Table 21: First Stage Results

Dependent Variable: $Sales_{j,t,t+4}$

(Dependent Variable of Second Stage: Log of investment)

	Over-identified Model			Just-identified Model	
	(1)	(2)	(3)	(4)	(5)
Aggregation level	firm	firm	sector	firm	sector
$Sales^{potential}_{j,t,t+4}$	0.672	0.672	0.343		
	[0.106]***	[0.447]	[0.416]		
$Sales^{actual}_{j,t-4,t}$	0.646	0.646	0.920	0.655	0.933
	[0.0366]***	[0.0650]***	[0.0731]***	[0.0636]***	[0.0673]***
Firm Controls	Yes	Yes	Yes	Yes	Yes
Std Errors/Clustering	Firm	Sector×Year	Robust	Sector×Year	Robust
No of Clusters	3819	95		95	
Observations	7,838	7,838	95	7,838	95
R-squared	0.966	0.966	0.992	0.966	0.992
First Stage F-Statistic	557.5	67.29	106.8	106.0	192.6

Notes: *** significant at 1 percent, ** significant at 5 percent, * significant at 10 percent. Standard errors are clustered on the indicated level for firm level regressions and robust standard errors shown for sector level regressions. All regressions include year fixed effects, industry fixed effects and firm controls of size, ownership, age, location and market competition. Observations in sector level regressions are averaged over firms in each sector and weighted by the number of firms in each sector-year cell. $Sales_{j,t,t+4}$ is defined as the average annualized change in ownership rates between t and $t+4$. The dependent variable in the second stage is the logarithm of investment.