INSTITUTE
OF ECONOMICS

Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM

Working Paper Series

The Wage-Productivity Nexus in the World Factory Economy

Giovanni Dosi [a]
Maria Enrica Virgillito [b,a]
Xiaodan Yu [a,c]

[a] Institute of Economics, Sant'Anna School of Advanced Studies, Pisa, Italy.
[b] Department of Economic Policy, Università Cattolica del Sacro Cuore, Milan, Italy.
[c] Nottingham University Business School China, University of Nottingham Ningbo, China.

2019/05 March 2019
ISSN(ONLINE) 2284-0400

The Wage-Productivity Nexus in the World Factory Economy

Giovanni Dosi[*†], Maria Enrica Virgillito[‡,†], and Xiaodan Yu[†,¶]

[†]Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy
[‡]Department of Economic Policy, Universita' Cattolica del Sacro Cuore, Milan, Italy
[¶]Nottingham University Business School China, University of Nottingham Ningbo, China

March 18, 2019

Abstract

This paper highlights new findings on the wage-productivity nexus in the World Factory Economy. After presenting the long-run macro-elasticity characterizing the phase of Chinese economic development since the eighties, we look at the wage-productivity nexus from a micro level perspective using a detailed firm-level dataset covering the period of ownership restructuring (1998-2007). A few results are quite robust under different estimation strategies. *First*, throughout the impressive Chinese economic miracle, elasticities of real wages to productivities – that is the ratios of rates of variations of the former to the latter – are always *positive* both under pooled and longitudinal estimates, both at firm- and sectoral-levels. *Second*, such elasticities are dramatically *low*, and falling in many distinct phases since the late seventies. That is, even in the manufacturing sector, the distribution of gains from the impressive labour productivity growth appears to be markedly uneven. Finally, *third*, governance institutions seem to matter a lot, with the majority of ownership types exhibiting firm-specific wage determination processes. The low elasticities of wages to productivity are plausibly the consequence of the massive flow of migrant workers from the rural areas to the coasts, somewhat resembling the early phase of the English Industrial Revolution with the pattern of enclosure in the country-side and massive migrations to the industrial towns.

JEL codes: L6, D22, D24, J31
Keywords: Wage, Productivity, Distributions, Chinese Industrialization

[*]Corresponding author: Institute of Economics, Scuola Superiore Sant'Anna, Piazza Martiri della Liberta' 33, I-56127, Pisa (Italy). E-mail address: `gdosi<at>santannapisa.it`

1. Introduction

This work studies the microeconomic dynamics of wages and productivity, and their relationship in Chinese manufacturing firms over the decade preceding the Great Recession. As known, increasing evidence suggests that in the industrialized West labour productivity has slowed down (Syverson, 2017), while its dispersion has increased (Dunne et al., 2004). As so did wages, which in some countries like U.S. remained nearly stagnant on average, but under increasing degrees of dispersion (Barth et al., 2016). In fact, in the "glorious decade" after WWII in the industrialized West, wages were fully indexed on productivity, often at sectoral or even national levels. However, more recently such indexation (so called degree of "pass-through") has generally fallen and moved to the firm-level (Schwellnus et al., 2017).

What happened in China? Here, against the foregoing Western tendencies, we shall examine the characteristics of the wage-productivity nexus in the World Factory Economy. While the patterns of productivity growth and convergence for the Chinese manufacturing sectors have been addressed in the literature (see, among others, Yu et al., 2015), less attention has been devoted to the analysis of the distribution of the latter impressive technological gains to wages. This is precisely what we shall do in the following. Using a detailed firm-level dataset which includes all industrial firms above some minimal scale threshold over the period 1998-2007, distinguishing also the ownership types of the firms (e.g., State-owned, foreign MNCs, private-owned, etc.), we investigate the inter-sectoral and inter-institutional heterogeneity in both wage and productivity distributions and the coupled dynamics of the two latter variables, both in level and growth rates.

First of all we document an overall process of convergence in both variables, even if more marked in productivity, and above all driven by the bottom part of the distributions (50-10 percentile ratio). When decomposing variance in terms of within and between sectoral variations, in both cases the within-sectoral (between-firms) component accounts for more than 80% of the total variation. In that, however, while the dispersion in productivity shrinks, the wage one remains nearly stable during the period under study.

In order to explore the link between wage and productivity we perform quantile regression estimates in levels and growth rates for the mean, as well as for the second moment. Interestingly, the relationship between wage and productivity substantially varies conditional on the ownership type, with Hong Kong/Macao/Taiwan-invested enterprises (HMTs), foreign-invested enterprises (FIEs), and domestic private owned enterprises (POEs), positively contributing to an increasing wage dispersion, while State-owned enterprises (SOEs) help in its reduction. However, what is remarkable are the low coefficients of elasticities which are always below 0.35, and often in the neighbourhood of 0.15. That is to say, most of the fruits of the impressive technological catching-up have not been passed through wages. When looking at growth rates, such elasticities vary across the quantiles of the wage growth distributions, according to a U-shaped relation, with the lowest and highest wage growth quantiles having relatively higher pass-through degrees. With respect to the wage-productivity variance relationship we find a positive monotonic link: hence sectors characterized by a less dispersed distribution in productivities display also relatively more equal wage distributions.

Finally, in order to exploit the time structure of the dataset while preserving the quantile regression approach, we also perform both a correlated random and a fixed effects quantile estimation. The estimated dynamic panel quantile regressions do confirm the different role played by SOEs vis-a'-

vis private ones, with only the former exerting an equalizing role. Overall, our results suggest the continuing coexistence of two processes of wage determination, distinct in terms of their degrees of idiosyncratic market responsiveness.

The paper is organized as follows: Section 2 briefly discusses the institutional transformation and some ensuing effects, while Section 3 presents the data structure together with some descriptive evidence on the coupled dynamics between wage and productivity. In Section 4 we perform a shift-and-share decomposition of wage and productivity dispersions to detect the different sources of heterogeneity. Section 5 addresses the firm-level link between wage and productivity by means of quantile regressions, both pooled and longitudinal. Finally, Section 6 provides some theoretical interpretations and concludes our analysis.

2. The institutional transformations: the broad picture

Let us start by placing the dynamics of wage and productivity in the broader context of the institutional and structural transformations which China underwent. During the period 1998 to 2007 there are at least three remarkable institutional changes that might have influenced the relationship between wage and productivity, namely the process of restructuring of SOEs, the adherence to the WTO, and finally the introduction of the minimum wage. Together, as analysed in Yu et al. (2015), China undertook an impressive process of catching-up, characterised by a dramatic growth in labour productivity. The latter was driven more by dynamics of *creative restructuring* of State-owned and State-participated firms rather than sheer Schumpeterian patterns of creative destruction. Indeed, the drivers of catching-up in China have been more the State-owned enterprises, and various forms of State-private ventures than the purely private ones.

After the Southern Tour of Deng Xiaoping in 1992, the process of restructuring of SOEs started, with an intensification phase from 1993 to 2003. The restructuring process was meant to render SOEs ever-more competitive in those sectors defined as strategic ones, such as telecommunication, computers, various "heavy industries", transportation and energy. Importantly, SOEs dramatically reduced their role as comprehensive welfare providers. As a result of this series of reforms, the occupational share of SOEs largely shrank, by shedding 28 millions of workers and reducing the number of SOEs from 120 to 32 thousands in 2004 (see Xia et al. 2014).

Dong (2005) investigates the dynamics of wage inequality and compares the drivers more linked to observable worker individual characteristics, such as education, with those linked to firm characteristics: the findings suggest that it is *where you work* and not *who you are* that more contributed to raising wage inequality. A complementary analysis regarding the role of SOEs in the evolution of wage dispersion is undertaken in Appleton et al. (2014) using CHIP urban household survey data. The authors document that since the beginning of the market transition of SOEs in 1986, although the centralised wage setting process was gradually dismantled, SOEs tended to more equally distribute bonuses to workers, in particular providing higher bonuses for low-wage workers, and relatively lower ones for high-wage workers. The opposite instead occurred inside private firms, wherein a more market-prone wage setting scheme has been adopted, rewarding more, the more proficient workers.

Another stream of literature has been looking at the relationship between wage inequality and trade openness. In particular, Han et al. (2012) document a pattern of increasing wage inequality, using Chinese Urban Household Survey data from 1988 to 2008, by means of a quantile regression strategy controlling for the impacts produced by the Southern Tour (1992) and the WTO China adherence (2001) for low, medium and high wage percentiles. Some other studies focus on the reverse causation, from wages to productivity, and look at the effects of minimum wage regulations, as China since 2004 has strongly reinforced the sanctions for not compliant firms. Hau et al. (2016) report on the so called cleansing effects of minimum wage. The higher labour costs might have triggered processes of internal restructuring.

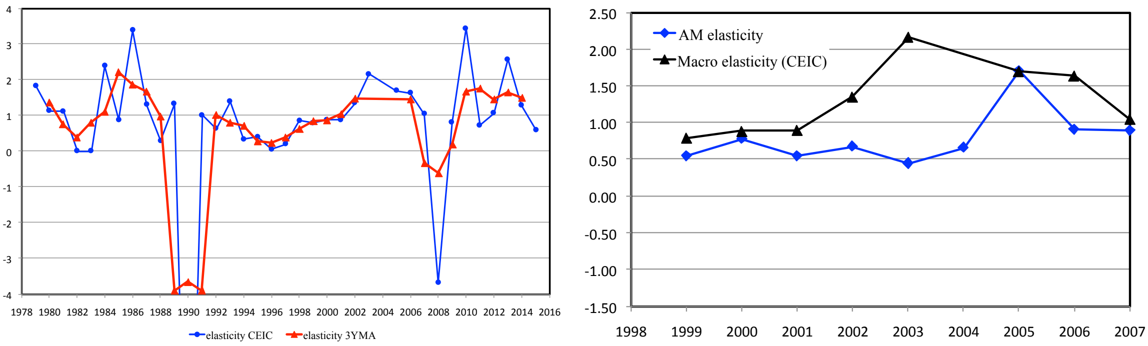
Overall the findings on wage convergence/divergence are rather controversial, with households data showing an increasing divergence, while manufacturing data, as we shall discuss, show convergence. In fact, households inequality grew notwithstanding the increase in nominal and real wage. On the one hand, some authors point at the potential erosion of the immense unlimited supply of labour from agricultural areas, as signalled by the "dramatic" nominal wage growth for migrant workers in the urban areas from 2003 to 2009 (Fang and Yang, 2011). On the other hand, as we shall show, labour productivity over the same period increased by almost one order of magnitude more: hence, whether the so called *Lewis turning point* has been reached is still a major question mark.

In the following we enrich the current understanding of the dynamics of wage dispersion and jointly study the dynamic of the pass-through from productivity to wages. In a complementary analysis, Card et al. (2018) show very low elasticity at micro level between wage and productivity dynamics. Basically, there are two major stylised facts which we are going to document in this work. First, elasticities are greater than zero – hence, strictly speaking, no "unlimited supply of labour" applies, and second, such elasticities are much smaller than one – indeed often around one-tenth – suggesting a pattern of income distribution biased toward increasing profits.

In order to provide a long run picture of the wage-productivity elasticity let us begin with aggregate manufacturing data. Figure 1, left panel, shows the wage-productivity elasticities, i.e. the ratio between the percentage change of real wage and the percentage change of labour productivity, for China's manufacturing sector since 1979 (both annual, blue line, and three years moving average, red line). One can observe three periods of decreasing elasticities (pass-through): the first one between mid-1980s to mid-1990s characterized by the decentralization of State power and an increasing decision-making autonomy of the SOEs; the second period between 2003-2008 (from the end of the restructuring of SOEs to the global financial crisis) characterized by the massive entry of domestic private-owned enterprises; the third one, from 2011 to the present, characterized by the slowing down of real wage growth.

Along this time-span we shall focus on the period covered by our micro-level dataset (see below). Note that Figure 1, left panel, is recovered by the World Bank dataset to estimate real value added and the CEIC dataset to construct the employment growth, while our dataset is restricted to firms above a certain turn-over threshold, covering the 90% percent of Chinese manufacturing value added. Among the latter, elasticities are systematically lower than in the former: cf. Figure 1, right panel. The discrepancy is due to a much lower productivity growth recovered by the World Bank dataset.

Figure 1: Left panel: wage-productivity elasticities in the manufacturing sector (1978-2016). Source: World Bank (real value added) and CEIC (employment growth). Right panel: comparison of elasticities, World Bank/CEIC vs. ASIE dataset (1998-2007). Source: NBSC, CEIC and the World Bank. Note: the time series is not continuous in 2004, because 2004 is a census year.



3. Data

3.1. Data description

We draw upon firm level data from the Annual Survey of Industrial Enterprise (ASIE) collected by the National Bureau of Statistics of China (NBSC). The dataset includes all industrial firms with sales above 5 million RMB covering the period 1998-2007 and has already been employed in other empirical investigations, among others, Yu et al. (2015).[1] The survey covers approximately 55 to 79 million workers, accounting for about 7.5% to 10.5% of the total Chinese employment.[2] Each firm is assigned to a sector according to the 4-digit Chinese Industry Classification (CIC) system that closely matches the Standard Industrial Classification (SIC) employed by the U.S. Bureau of Census.[3] Out of the comprehensive set of all firms, we focus on manufacturing firms only (CIC 13 - 42): Table 1 shows their summary statistics. The total number of employees in the manufacturing sector has increased from 50 in 1998 to 68 million in 2007, after a fall by 5.7 million from 1998 to 2001. The total output has increased from 5.93 to 35 trillion RMB in the same period and the number of firms from 14 to 30 thousand units approximately (see Table A.2 for the number of firms by ownership type). In the analysis that follows we apply a few cleaning procedures in order to eliminate visible recording errors, yielding what we call "China Micro Manufacturing" (CMM).[4] And we keep firms existing for at least two consecutive years.

3.2. Variables

Productivity (π_{ijt}) is the ratio of value added (at 1998 constant prices) over the number of employees, in logs. It is deflated by the 4-digit output deflator (source: Brandt et al. 2012). Firm's total labour compensation is composed by wages, unemployment insurance and welfare benefits. Wage (w_{ijt}) is

[1]Industry if defined to include mining, manufacturing and public utilities, according to NBSC. Five million RMB is approximately $US 600,000.

[2]The number of employees by firm ownership type is presented in Table A.3.

[3]In 2003, the classification system was revised. Some sectors were further disaggregated, while others were merged together. To make the industry code comparable over time, we adopted the harmonized classification proposed in Brandt et al. (2012).

[4]We dropped firms with missing, zero or negative output, value-added, sales, original value of fixed assets, with employment < 8.

Table 1: Summary statistics (total) of the Chinese manufacturing firm-level dataset.

Year	Number of Firms	Value Added	Sales	Output	Emp	Wage	Welfare	Cost of labour
1998	148 661	1.52	5.48	5.94	50.72	342.93	46.02	0.44
1999	146 075	1.68	5.96	6.37	47.36	351.00	46.18	0.45
2000	147 246	1.96	7.14	7.48	45.83	387.10	50.98	0.49
2001	155 659	2.22	7.99	8.40	44.95	416.39	54.01	0.50
2002	165 793	2.62	9.37	9.79	45.87	471.95	58.29	0.56
2003	181 001	3.40	12.38	12.72	48.71	549.46	67.81	0.65
2004	258 869	4.80	17.14	17.74	56.52	725.44	81.29	0.84
2005	250 952	5.71	21.34	21.74	59.21	885.13	101.51	1.02
2006	278 644	7.23	26.99	27.40	63.32	1090.65	123.65	1.25
2007	312 284	9.37	34.70	35.27	68.38	1415.58	139.62	1.60

Note: all values are denoted in trillion RMB (wage and welfare are in billion RMB) and employment in millions of workers. All manufacturing firms are included. The output of year 2004 is not directly available from the original dataset, thus, we proxy it using "sales - year beginning inventory + year end inventory + value added tax".

Table 2: Summary statistics (mean) of the dataset after cleaning.

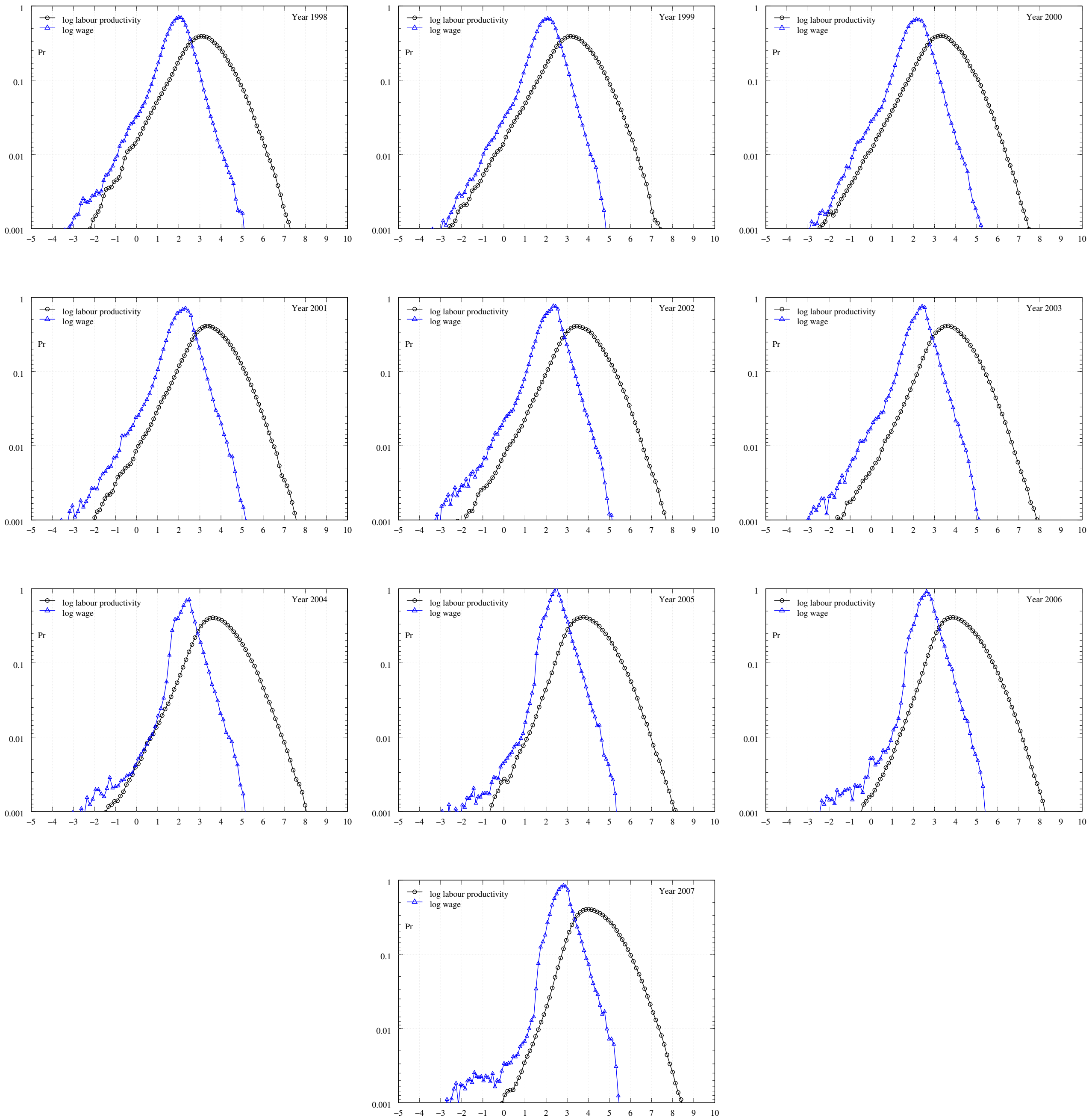
Year	Number of Firms	Labour Productivity	Wage	(log-) Labour Productivity	(log-) Wage	Growth of Labour Productivity	Growth of Wage
1998	108286	44	9	3.101	1.873	NA	NA
1999	125917	48	9	3.181	1.936	0.070	0.076
2000	126054	54	10	3.314	2.038	0.061	0.076
2001	138410	59	10	3.439	2.071	0.046	−0.008
2002	149189	68	11	3.569	2.152	0.083	0.067
2003	162086	76	12	3.716	2.230	0.099	0.063
2004	211534	88	13	3.817	2.345	0.047	0.089
2005	238160	97	14	3.957	2.476	0.154	0.155
2006	265912	114	17	4.118	2.616	0.171	0.158
2007	248299	137	19	4.315	2.764	0.177	0.140

Note: labour productivity and wages are at 1998 constant price, in 1000 RMB. Growth rates are calculated as log differences of real values over two consecutive years. Source: our elaboration on CMM.

the ratio of firm's total labour compensation (at 1998 constant prices) over the number of employees, in logs. It is deflated by consumer price index (source: National Bureau of Statistics of China). Table 2 shows the summary statistics. Figure 2 shows the kernel density distributions of wages and labour productivities. Already at a first glance some interesting patterns do emerge. While, not surprisingly, the support of the productivity distribution only partly overlaps with the wage distribution, overtime the support of the former clearly moves to the right, but the support of the latter remains roughly constant. Together, a *lower* mode in the wage distribution hints at an increasing wage-productivity gap for the right tail of the distribution itself. Figure 3 shows the kernel density distributions of the growth rates of wage and labour productivity. The two distributions roughly overlap until 2001. Interestingly, both the lower and the upper tails of the wage growth distribution become much fatter since 2002, after the adherence to the WTO, indicating an increasing granularity of wage growth rate.

We identify seven categories of firms according to their ownership and governance structures. They are State-owned enterprises (SOEs); collective-owned enterprises (COEs), Hong Kong, Macao and Taiwan-invested enterprises (HMTs); foreign-invested enterprises (FIEs), including foreign MNCs

Figure 2: Distribution of (log) wages and labour productivities (at 1998 constant price), years 1998-2007. Pooling all firms in manufacturing. Source: our elaboration on CMM.



(FMNC) and joint ventures (JV) with a foreign share above 25%; shareholding enterprises (SHEs), that is State-private Chinese joint ventures; private-owned enterprises (POEs); and other domestic enterprises (ODEs).[5] Figure 4 and Figure 5 show the evolution of the means of wage and labour productivities, *levels and growth rates* respectively, by six major ownership types. In Figure 4, we observe an increasing gap between productivity level and wage level for all ownership types. Figure 6 and Table 3 show the evolution of the wage-productivity elasticities by six major ownership types. Again,

[5]As reported in Table A.1 in the Appendix, the original 23 registration categories have been aggregated in line with Jefferson et al. (2003).

Figure 3: Distribution of growth rates of wage and growth rates of labour productivities, year 1999-2007. Pooling all firms in manufacturing. Source: our elaboration on CMM.

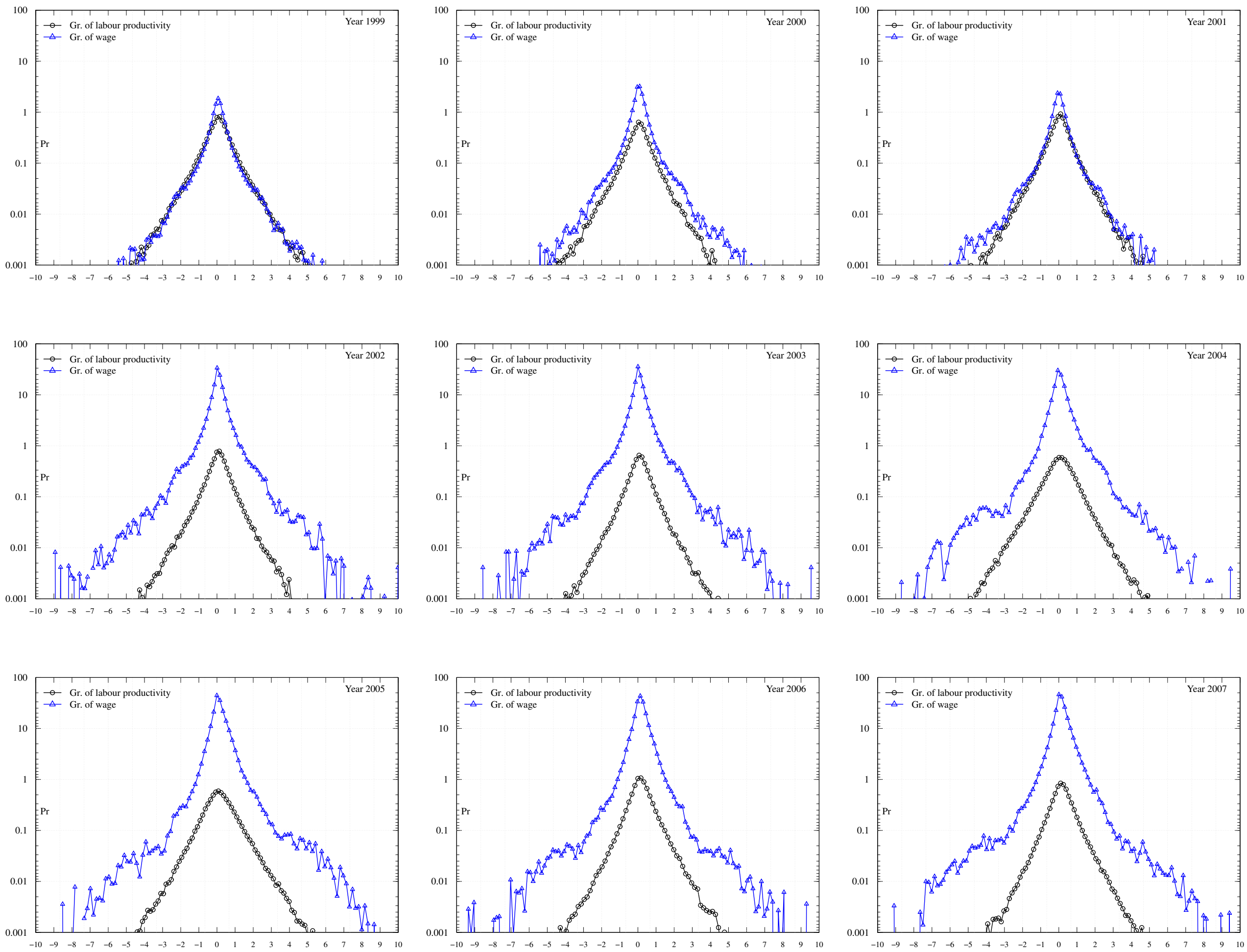


Figure 4: By six major ownership types: means of wages and labour productivities levels. Source: our elaboration on CMM.

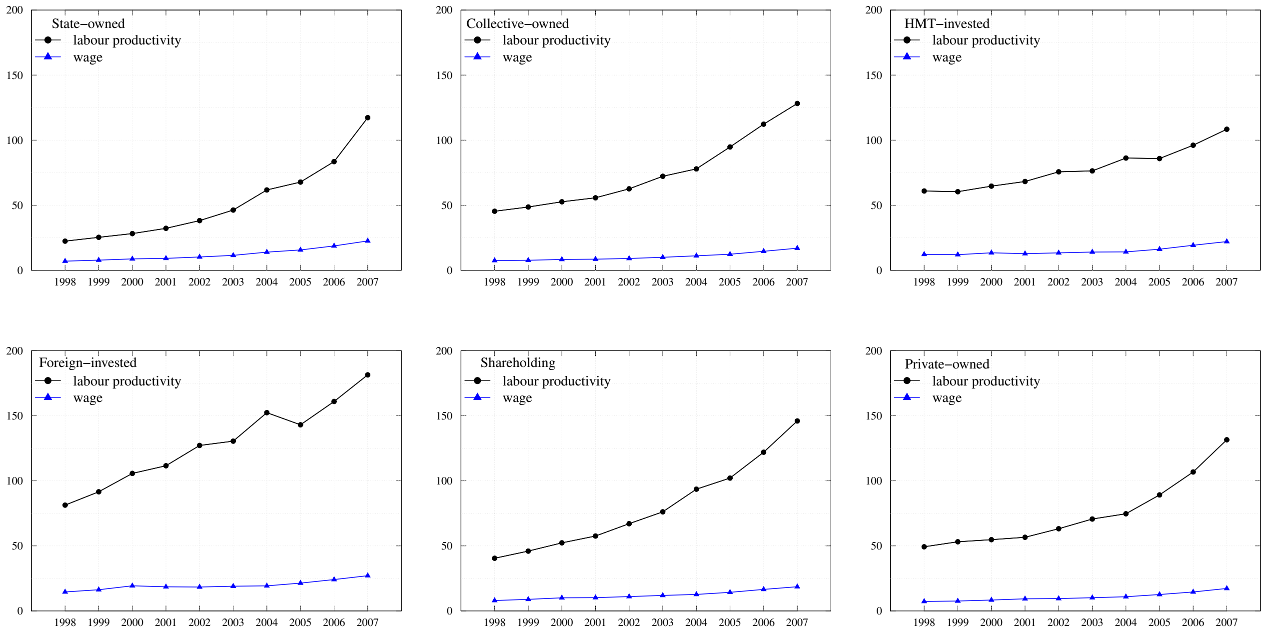


Table 3: Elasticities by ownership types. Source: Figure 5

Ownership types	1999	2000	2001	2002	2003	2004	2005	2006	2007	1999-2007
State-owned	1.70	1.50	-1.68	2.16	0.67	1.11	3.05	1.11	0.51	1.14
Collective-owned	0.96	1.41	0.06	0.72	0.61	26.20	0.61	1.23	1.15	0.92
HMT-invested	1.25	1.67	-0.36	0.63	3.43	0.67	5.07	1.28	1.09	1.40
Foreign-invested	0.99	0.81	-1.67	1.57	0.54	0.48	4.45	0.97	0.95	1.05
Shareholding	0.95	1.05	0.07	0.74	0.66	2.06	0.92	1.01	0.63	0.84
Private-owned	0.66	0.97	0.38	0.66	0.49	6.83	0.73	0.80	0.78	0.80

the year of WTO adherence is associated with a major negative shock in the latter elasticities due to a dramatic decline in the rate of wage growth which in some of the ownership types becomes even negative. SOEs are among those two reporting the highest ratio in the entire time span.

4. Convergence in wages and productivities dispersions across firms

Moving a step deeper into the analysis of the degrees of heterogeneity, let us present alternative measures of dispersion for productivity and wages within narrowly defined sectors and/or ownership types.

4.1. 90-10 ratio

The $90-10$ wage (productivity) ratio is defined as the ratio of the $90th$ percentile to the $10th$ percentile of the wage (productivity) distribution. Figure 7 shows the converging trend of both between-firm wage and productivity dispersions. The average wage in the highest paying firms, i.e. those at the $90th$ percentile of the wage distribution, were 2.7 times those at the bottom decile in 1998, and the

Figure 5: By six major ownership types: means of the growth rates of wages and labour productivities. Source: our elaboration on CMM.

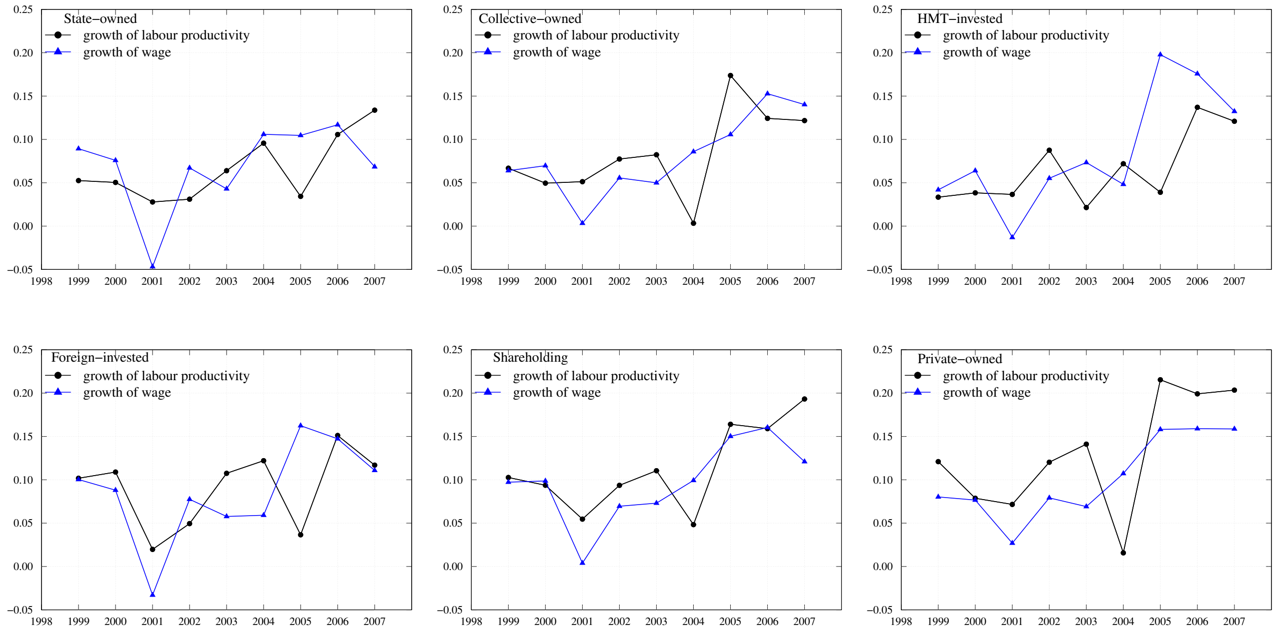
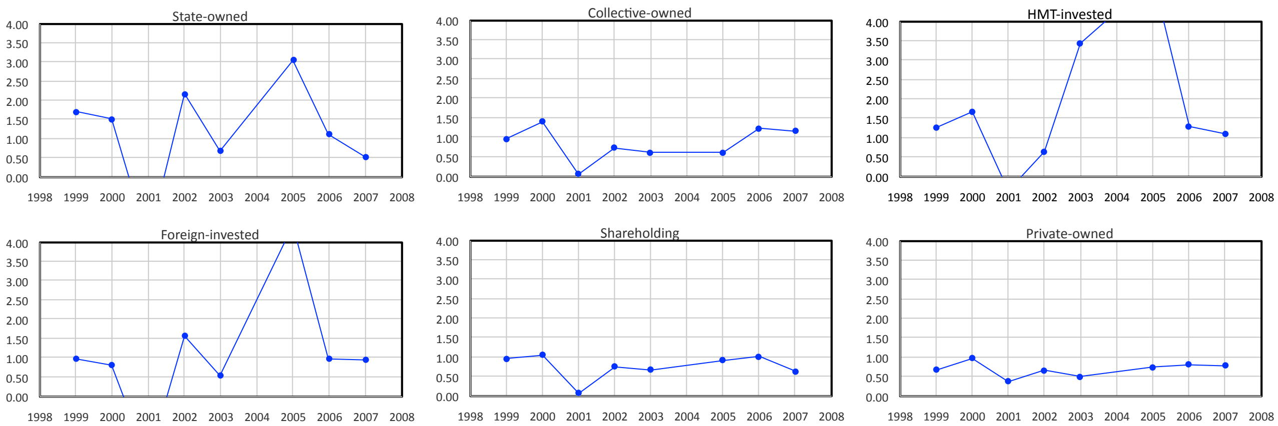


Figure 6: By six major ownership types: elasticities (growth of real wage per employee over growth of labour productivity. Source: Figure 5.



ratio decreased to 1.7 in 2007. The productivity of firms at the $90th$ percentile of the productivity distribution were 2.7 times higher than those at the bottom decile, and 1.8 times in 2007. This sharp decline occurred for all ownership structures as shown in Figure 8 and Table 4. Interestingly, SOEs present the highest initial ratios for wage and productivity dispersion in 1998, but they display also the steepest fall compared to other ownership types.

Figure 7: Wage/productivity 90-10 ratio by year, all manufacturing firms. [Note: equal weights.] Source: our elaboration on CMM.

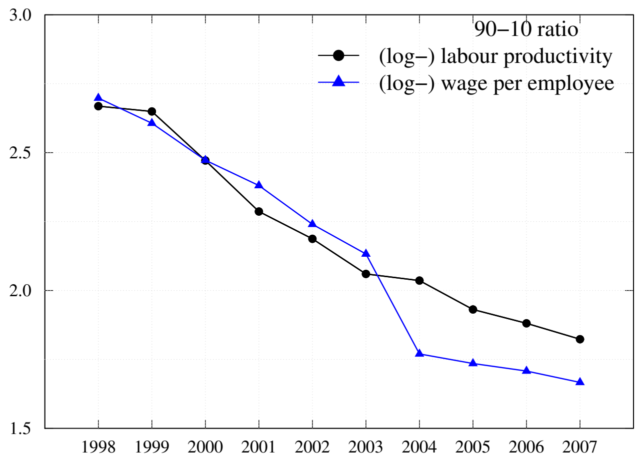


Figure 8: By six major ownership types: wage/productivity 90-10 ratio by year, all manufacturing firms. [Note: equal weights.] Source: our elaboration on CMM.

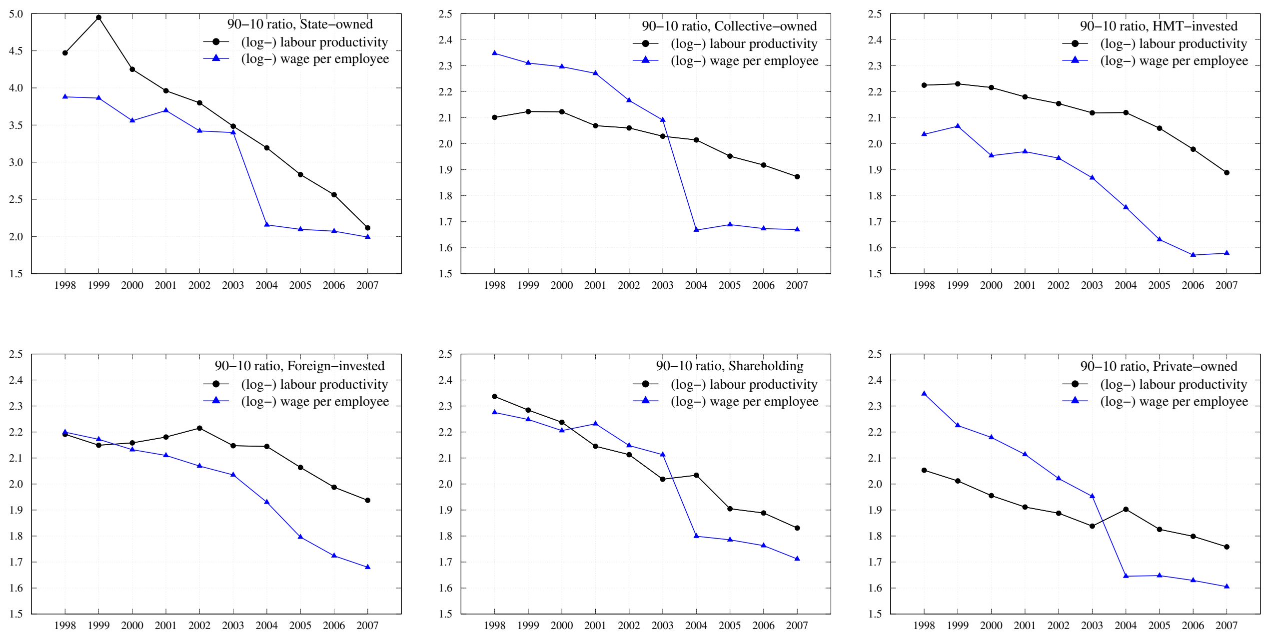
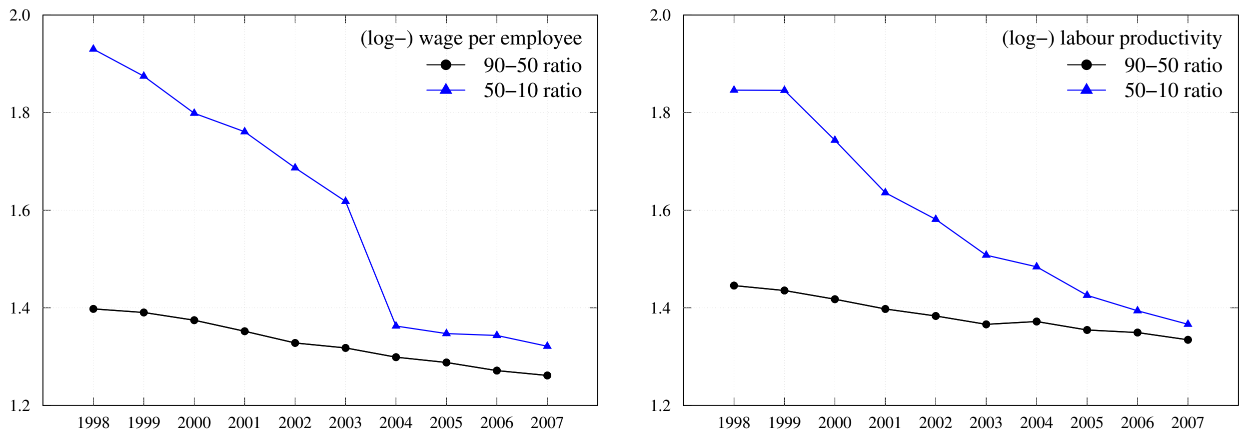


Table 4: By six major ownership types: wage/productivity 90-10 ratios, 90-50 ratios and 50-10 ratios.

	90-10 ratio				90-50 ratio				50-10 ratio			
	Productivity		Wage		Productivity		Wage		Productivity		Wage	
Ownership	1998	2007	1998	2007	1998	2007	1998	2007	1998	2007	1998	2007
State-owned	4.5	2.1	3.9	2.0	1.6	1.4	1.5	1.3	2.9	1.6	2.7	1.5
Collective-owned	2.1	1.9	2.3	1.7	1.4	1.3	1.3	1.2	1.5	1.4	1.8	1.3
HMT-invested	2.2	1.9	2.0	1.6	1.4	1.4	1.3	1.2	1.6	1.4	1.5	1.3
Foreign-invested	2.2	1.9	2.2	1.7	1.4	1.4	1.4	1.3	1.6	1.4	1.6	1.3
Shareholding	2.3	1.8	2.3	1.7	1.4	1.3	1.3	1.3	1.7	1.4	1.7	1.3
Private-owned	2.1	1.8	2.3	1.6	1.4	1.3	1.3	1.2	1.5	1.3	1.8	1.3

Figure 9: Wage/productivity at the top (90-50 ratio) versus bottom (50-10 ratio) of the distribution by year, all manufacturing firms. [Note: equal weights.] Source: our elaboration on CMM.



4.2. 90-50 and 50-10 ratios

In order to understand the source of the convergence, we split the $90-10$ ratio in two components. The $90-50$ wage (productivity) ratio is defined as the ratio of the $90th$ percentile to the $50th$ percentile (the median) of the wage (productivity) distribution. It captures dispersion in the upper tail of the distribution. Symmetrically, the $50-10$ wage (productivity) ratio is the ratio of the $50th$ percentile to the $10th$ percentile of the wage (productivity) distribution, capturing the dispersion in the bottom tail of the distribution.

The evidence suggests that convergence is mainly driven by the push at the bottom of the wage distribution while only a very mild convergence has happened at the top (Figure 9 [left]). Interestingly there is a significant drop of the degree of divergence in the bottom part between 2003 and 2004, probably due also to the enforcement of the minimum wage legislation in 2004. Figure 9 [right] shows the evolution of the top and the bottom of the productivity distributions, suggesting, again, that convergence has been more at the bottom, starting around 1999, when the process of SOEs restructuring was almost completed.

Disaggregating by ownership types, Figure 10 shows the evolution of the top and bottom of the wage distribution. The discrete drop between 2003 and 2004 is more pronounced for all domestic ownership types, while wage adjustments have been smoother in foreign-invested firms. Figure 11 reports the same analysis for productivity. Interestingly, the strongest converging trend at the *top* of productivity distribution concerns the domestic privately enterprises (POEs) and shareholding ones (SHEs). Note that convergence at the top might mean two opposite things, namely, that the top slows down

Figure 10: By six major ownership types: wage at the top (90-50 ratio) versus bottom (50-10 ratio) of the distribution by year, all manufacturing firms. [Note: equal weights.] Source: our elaboration on CMM.

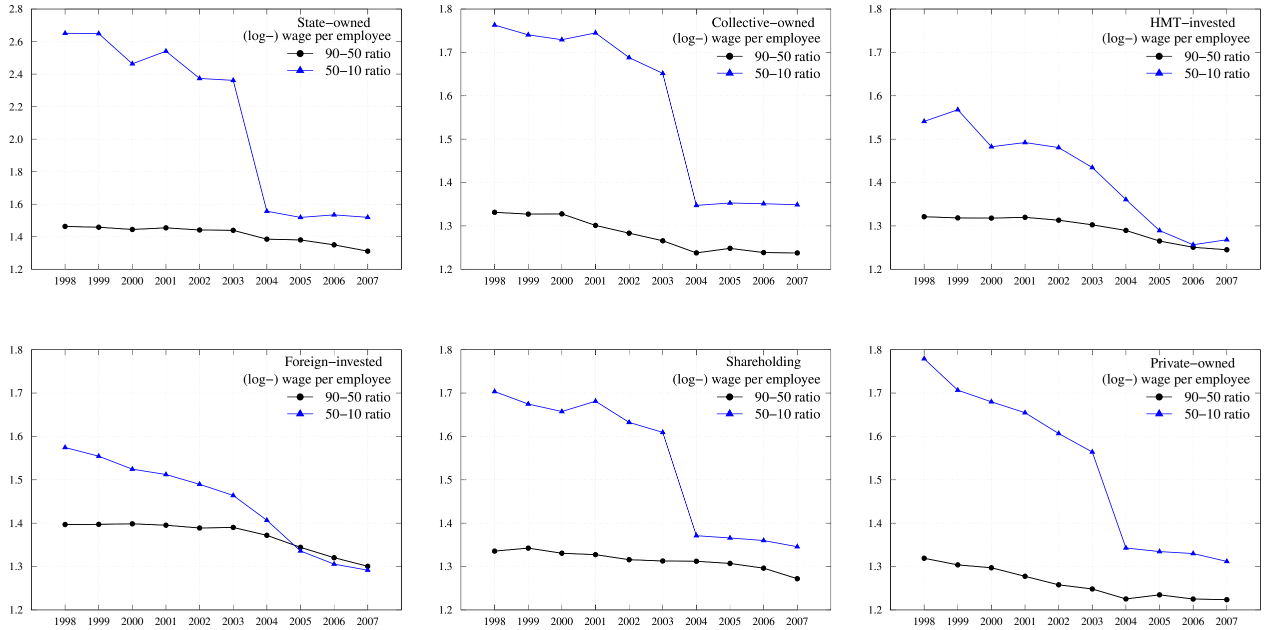
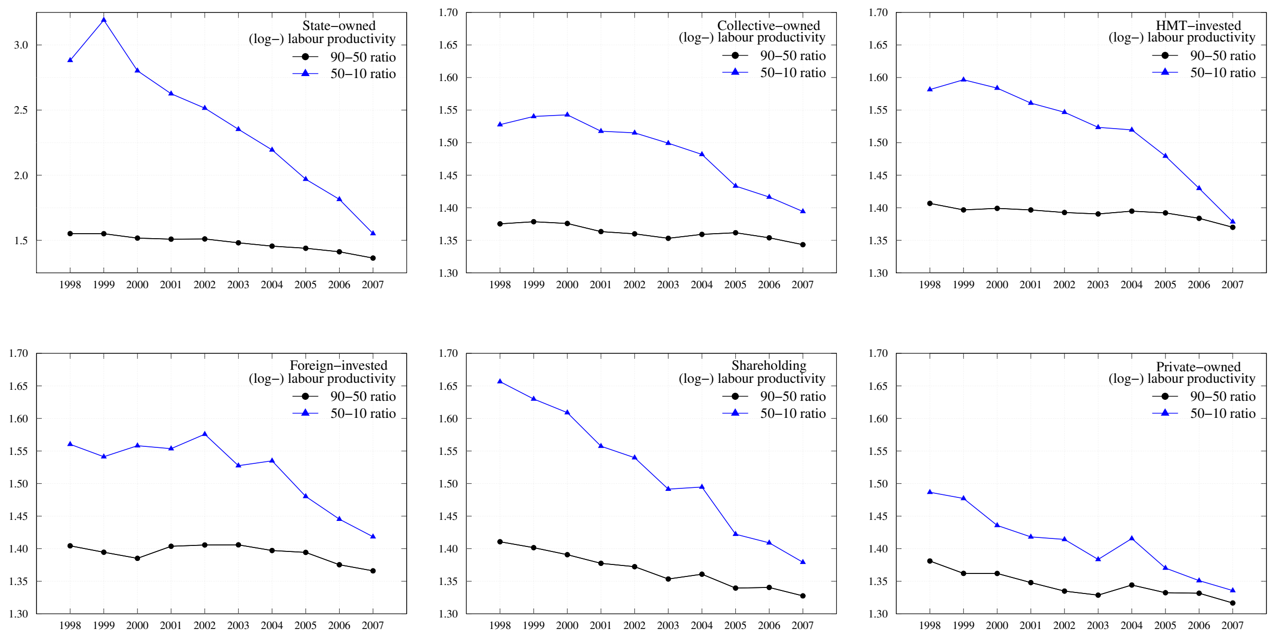


Figure 11: By six major ownership types: productivity at the top (90-50 ratio) versus bottom (50-10 ratio) of the distribution by year, all manufacturing firms. [Note: equal weights.] Source: our elaboration on CMM.



with respect to the median, or, conversely, that the median catches-up with the top. Conversely, the strongest converging trend at the *bottom* concerns state-owned enterprises (SOEs) and shareholding ones (SHEs).

4.3. Decomposition of wage/productivity variance

Having detected some convergence processes in the distribution of both wages and productivity driven by the bottom part of the distribution, let us investigate by means of a shift and share decomposition the relative contribution of the *within* and *between* sectoral components. The (labour-weighted) variance of wage ($Var\ w_{ijt}$)/ productivity ($Var\ \pi_{ijt}$) is decomposed into a within-sector component and a between-sector one, according to:

$$\underbrace{\sum_i \frac{L_{ijt}}{L_t}(w_{ijt}-\overline{w}_t)^2}_{\text{Var } w_{ijt}} \equiv \underbrace{\sum_j \frac{L_{jt}}{L_t}\sum_{i\in j}\frac{L_{ijt}}{L_{jt}}(w_{ijt}-\overline{w}_{jt})^2}_{\text{within}} + \underbrace{\sum_j \frac{L_{jt}}{L_t}(\overline{w}_{jt}-\overline{w}_t)^2}_{\text{between}} \tag{1}$$

where L_{ijt} is the number of employees of firm i in sector j at time t; L_{jt} is the total number of employees of sector j at time t; L_t is the total number of employees at time t; $\overline{w}_t = \sum_i\sum_j \frac{L_{ijt}}{L_t}w_{ijt}$ is the grand (labour) weighted mean of wages; $\overline{w}_{jt} = \sum_{i\in j}\frac{L_{ijt}}{L_{jt}}w_{ijt}$ is the sectoral (labour) weighted mean of wages. The decomposition is done by pooling all firms in manufacturing for each 2-digit sector.[6]

As shown in Table 5 and in Table 6, the within-sector (between-firms) component accounts for around 80% of the wage (productivity) dispersion in each 2-digit sector, every year. The results corroborate an increasing empirical literature documenting that the *between firms* wage/productivity variation is the major driver of the observed heterogeneity. Figure 12 shows the evolution of the shares of the within-sectoral wage/labour productivity dispersions: while the within sectoral component of productivity dispersion presents a clear decreasing trend, the same does not occur for the sectoral component of wage dispersion, which oscillates between 87% and 84% in the period under study. It is an interesting and suggestive piece of evidence. The fall of the within-sector component in productivity dynamics hints at a generalized catching-up process cutting across all manufacturing, steadily reducing dualism, as the classics of development theory would have argued, between a modern/dynamic part of the industry and an informal/backward one. Conversely, the persistency of the within component in wages strongly suggests a persistent heterogeneity in labour market conditions, and in bargaining and power relations between firms and workers.

The conjecture is supported by the evidence on the evolution of the share of within-sector wage and labour productivity dispersions by the six major ownership types (see Figure 13). This allows to disentangle the different behaviour of State owned vs. private firms. So the within share of the total wage variance of SOEs is much lower than other ownership types and decreasing from 78% in 1998 to 60% in 2007. Conversely, that of domestic private firms (POEs) increases from 80% to 95% in 2007. A similar pattern does also characterize the within share of productivity variance, with a decreasing within share of SOEs, from 70% to 50% in 2007 and an increasing within share of POEs, from 70% to 85%. That is, under our foregoing conjecture, SOEs converged more consistently as a whole.

[6]Note: deviation from the labour-weighted 4-digit wage mean.

Table 5: Share of within-sector wage variance. [Note: variance is labour weighted.]

CIC	SECTOR	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
	Whole manufacturing	87.0	85.3	85.5	85.6	84.2	84.9	84.1	85.1	85.9	87.1
13	Processing of food from agricultural products	92.2	94.4	95.1	87.8	85.4	94.9	97.5	97.3	98.4	98.3
14	Foodstuff	93.2	96.2	94.6	94.3	90.3	94.8	94.8	94.3	94.5	95.8
15	Manuf. of beverages	92.9	91.7	91.9	93.7	92.3	93.2	86.3	90.0	90.3	89.9
16	Manuf. of tobacco	89.2	89.4	88.3	90.6	87.0	82.3	61.7	80.1	85.1	66.8
17	Manuf. of textile	97.6	96.4	95.5	96.6	96.0	96.5	96.3	96.7	95.1	96.3
18	Manuf. of textile wearing apparel, footwear, cand caps	98.8	99.3	99.7	99.7	99.0	99.5	99.7	99.9	99.8	99.5
19	Manuf. of leather, fur, feather and related products	99.0	98.1	98.5	98.5	99.0	99.1	99.6	99.5	99.6	98.7
20	Processing of timber, manufacture of wood, bamboo,	97.3	96.9	95.9	95.3	97.5	94.7	97.3	96.1	97.4	97.0
21	Manuf. of furniture	98.8	98.2	96.8	95.2	96.5	97.8	98.7	97.9	97.0	98.2
22	Manuf. of paper and paper products	99.5	99.4	99.2	99.2	99.6	99.3	99.3	99.1	99.3	98.0
23	Printing, reproduction of recording media	98.9	99.2	98.1	97.7	98.5	99.1	99.5	99.7	98.8	99.5
24	Manuf. of articles for culture, education and sport activity	94.5	93.8	95.3	95.8	95.3	96.9	96.4	97.8	98.7	98.7
25	Processing of petroleum, coking, processing of nuclear fuel	89.1	77.9	80.8	93.5	79.3	73.0	76.3	71.0	72.0	82.3
26	Manuf. of raw chemical materials and chemical products	93.0	75.8	93.4	92.6	93.8	90.4	91.5	92.9	93.2	93.8
27	Manuf. of medicines	97.9	98.1	97.5	98.7	98.3	96.7	97.2	97.1	98.0	98.4
28	Manuf. of chemical fibers	93.0	85.2	88.2	91.3	93.6	94.6	92.7	75.7	83.3	85.2
29	Manuf. of rubber	98.3	94.1	88.1	95.6	94.4	93.9	94.2	93.0	95.5	96.1
30	Manuf. of plastics	97.4	97.2	96.1	96.0	98.4	97.2	96.9	97.1	95.9	96.5
31	Manuf. of non-metallic mineral products	96.4	95.2	94.1	94.5	94.6	95.6	95.6	92.8	93.9	94.4
32	Smelting and pressing of ferrous metals	84.9	84.5	82.7	82.9	77.1	82.0	85.7	85.7	88.0	88.8
33	Smelting and pressing of non-ferrous metals	87.2	91.1	84.4	88.9	83.9	86.6	90.0	88.2	88.9	90.3
34	Manuf. of metal products	93.4	94.3	93.9	98.0	93.7	95.8	94.8	95.8	97.6	96.7
35	Manuf. of general purpose machinery	92.2	93.8	92.2	96.1	95.3	94.2	94.5	95.4	94.3	94.7
36	Manuf. of special purpose machinery	92.1	90.7	89.2	89.5	89.5	88.8	92.0	93.0	92.3	90.3
37	Manuf. of transport equipment	89.9	91.7	86.4	86.6	84.7	84.5	85.7	87.7	82.7	88.8
39	Manuf. of electrical machinery and equipment	95.0	90.0	97.4	96.2	95.9	95.0	93.6	95.8	95.9	96.2
40	Manuf. of communication equipment, computers and other electronic equipment	90.7	89.3	89.2	91.7	88.4	88.7	77.9	84.2	89.5	88.5
41	Manuf. of measuring instruments and machinery for cultural activity and office work	89.7	93.3	85.9	95.4	92.4	93.2	91.6	90.6	86.9	88.5
42	Manuf. of artwork and other manufacturing	83.4	90.4	89.9	89.7	96.2	93.3	95.3	95.5	94.0	95.1

Table 6: Share of within-sector labour productivity variance. [Note: variance is labour weighted.]

CIC	SECTOR	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
	Whole manufacturing	82.7	83.5	83.5	82.4	81.1	78.7	78.0	78.7	78.0	77.7
13	Processing of food from agricultural products	89.6	93.9	95.1	93.3	91.9	92.8	92.4	92.0	92.1	90.3
14	Foodstuff	83.9	89.4	86.6	86.8	86.8	89.0	92.2	90.4	89.5	90.0
15	Manuf. of beverages	92.6	94.4	93.9	93.4	92.2	90.8	91.8	95.0	95.6	95.3
16	Manuf. of tobacco	85.3	84.1	84.5	86.5	88.6	73.6	50.3	61.6	52.6	51.7
17	Manuf. of textile	94.0	94.5	94.5	93.1	92.6	94.4	96.5	95.9	96.1	96.8
18	Manuf. of textile wearing apparel, footwear, cand caps	99.2	98.2	97.2	98.7	99.3	98.6	99.3	99.3	98.9	99.5
19	Manuf. of leather, fur, feather and related products	96.1	96.0	93.2	91.9	91.5	86.1	88.0	85.1	86.4	83.7
20	Processing of timber, manufacture of wood, bamboo,	98.5	96.6	97.4	97.4	95.8	90.8	94.0	93.4	94.6	93.6
21	Manuf. of furniture	97.0	97.7	97.8	97.2	97.1	98.9	96.6	97.4	98.2	98.7
22	Manuf. of paper and paper products	99.3	99.7	99.3	99.8	99.3	99.3	98.8	97.7	97.6	96.8
23	Printing, reproduction of recording media	94.1	96.2	94.6	91.5	94.0	93.2	95.6	96.2	96.0	95.7
24	Manuf. of articles for culture, education and sport activity	92.4	90.7	89.5	86.5	91.3	90.1	92.7	93.5	93.1	93.4
25	Processing of petroleum, coking, processing of nuclear fuel	72.0	68.8	81.9	71.7	69.3	69.2	67.4	76.4	75.6	82.5
26	Manuf. of raw chemical materials and chemical products	90.6	81.6	89.9	89.4	90.2	86.5	85.8	90.2	91.2	91.4
27	Manuf. of medicines	97.1	97.3	97.0	96.6	97.1	95.5	97.0	96.2	96.9	96.6
28	Manuf. of chemical fibers	91.8	88.4	85.4	92.1	96.1	95.4	98.9	87.3	82.7	81.4
29	Manuf. of rubber	79.3	83.9	80.9	82.2	82.9	79.5	79.9	76.6	77.8	76.6
30	Manuf. of plastics	97.4	95.2	95.2	95.0	93.4	94.3	92.0	92.2	92.0	90.6
31	Manuf. of non-metallic mineral products	93.2	92.7	90.6	88.7	86.7	87.2	87.0	86.1	88.0	87.4
32	Smelting and pressing of ferrous metals	94.8	94.0	93.1	91.9	89.2	89.9	91.5	90.1	90.8	94.8
33	Smelting and pressing of non-ferrous metals	96.5	94.1	96.6	92.8	92.2	92.5	92.7	90.0	89.6	90.0
34	Manuf. of metal products	93.0	94.3	94.3	94.1	94.5	94.0	89.9	92.6	93.0	93.4
35	Manuf. of general purpose machinery	93.8	96.5	95.1	95.5	95.7	93.4	94.5	93.7	92.2	93.0
36	Manuf. of special purpose machinery	85.2	86.6	80.1	85.6	87.5	86.6	90.7	94.5	93.7	92.5
37	Manuf. of transport equipment	89.0	89.9	92.7	87.4	79.7	75.7	83.6	81.1	79.0	80.1
39	Manuf. of electrical machinery and equipment	85.0	85.4	90.5	86.7	89.0	88.0	86.7	87.8	86.5	84.9
40	Manuf. of communication equipment, computers and other electronic equipment	84.8	87.3	84.9	86.0	84.4	86.4	86.9	88.7	88.4	89.8
41	Manuf. of measuring instruments and machinery for cultural activity and office work	88.4	86.4	88.6	89.5	91.4	87.0	84.7	83.5	78.6	76.9
42	Manuf. of artwork and other manufacturing	83.1	94.1	94.9	94.1	95.8	96.7	96.3	95.9	96.2	97.2

Figure 12: The share of within-sector wage/labour productivity dispersion. [Note: pooling all firms in manufacturing sector, deviation from the 4-digit sectoral labour-weighted mean.] Source: our elaboration on CMM.

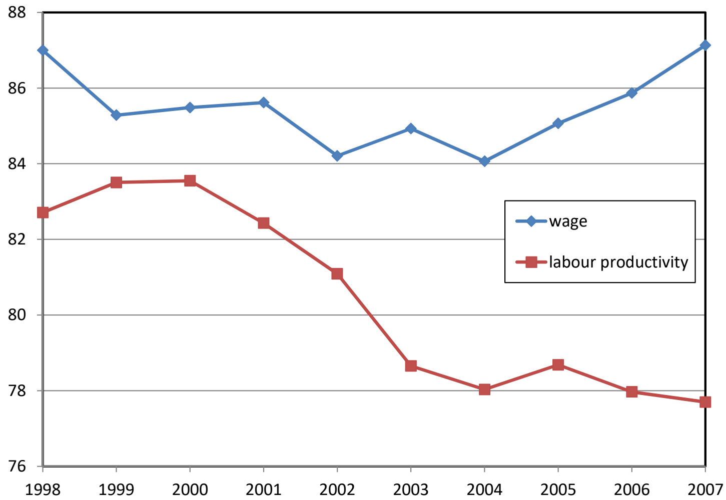
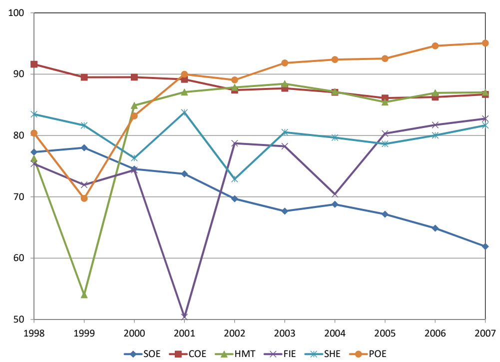
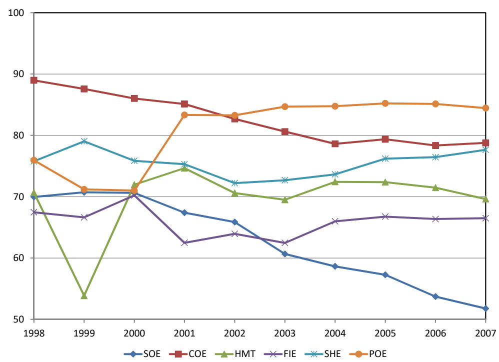


Figure 13: By ownership types: the share of within-sector wage [left] and labour productivity [right] dispersion. [Note: pooling all firms in manufacturing sector, deviation from the 4-digit sectoral labour-weighted mean.] Source: our elaboration on CMM.

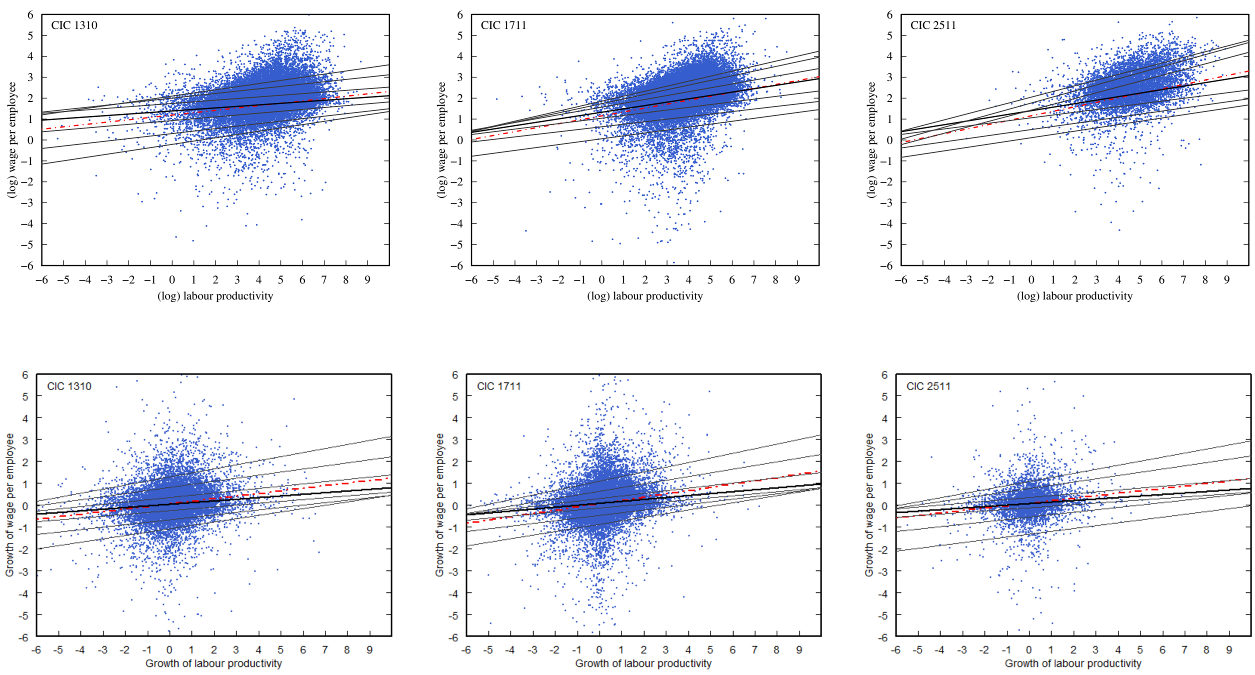




5. A quantile regression analysis of the wage-productivity nexus

In order to further analyse the wage-productivity nexus, let us first consider standard pooled OLS estimates. Recall from Figure 2 that looking at levels, both wage and productivity distributions display skewness and fat tails. At the same time growth rates are tent-shaped. All this evidence on deep heterogeneity and lack of Gaussian features, militates in favour of the use of quantile regression analysis. With respect to our own dataset, Figure 14 presents the scatter plot of the wage-productivity nexus, both in level (top panel) and in growth rates (bottom panel), for three representative 4-digit sectors. The dashed lines present the least-square estimates, while the grey lines do present the quantiles in the range of $0.05 - 0.95$. If the least-squares estimate had correctly captured the relationship, all the grey lines and the dashed line would have been parallel. On the contrary already at a first glance does emerge an increasing dispersion for higher and lower quantiles vis-a'-vis the median one. This prompts the use of quantile regression techniques robust as they are to outliers and heavy tailed distributions (Koenker and Hallock, 2001).

Figure 14: Scatterplot and quantile regression fit. The plots in the first row show a scatter plot on wage level vs. productivity level of three representative 4-digit sectors (manufacturing of textile clothing, plastic, communication transmitting equipment). The plots in the second row show a scatter plot on wage growth vs. productivity growth. Superimposed on the plots are the 0.05, 0.1, 0.25, 0.75, 0.90, 0.95 quantile regression lines in grey, the median fit in solid black, and the least-squares estimate of the conditional mean function as the dashed line. Source: our elaboration on CMM.



5.1. Quantile regression

The quantile regression model (Koenker and Bassett Jr, 1978) reads as:

$$y_{it} = x'_{it}\beta_\tau + u_{\tau it} \quad \text{with} \quad Q_\tau(y_{it}|x_{it}) = x'_{it}\beta_\tau \tag{2}$$

where y_{it} is the dependent variable, x is a set of regressors, β is the vector of parameters to be estimated, and u is a vector of residuals. $Q_\tau(y_{it}|x_{it})$ stands for the τ^{th} conditional quantile of y_{it} given x_{it}. The τ^{th} conditional quantile solves the minimization problem:

$$\beta_\tau \equiv \underset{b}{\text{argmin}}\ E[\rho_\tau(y_{it} - x'_{it}b)] \tag{3}$$

where $\rho(u) = 1(u > 0) \cdot \tau|u| + 1(u \leq 0) \cdot (1 - \tau)|u|$ is called the "check function". If $\tau = 0.5$, this turns out in terms of least absolute deviations. In this case, $Q_\tau(y_{it}|x_{it})$ is the conditional median since the conditional median minimizes absolute deviations. Otherwise, the check function weights positive and negative terms asymmetrically. The quantile regression estimator, $\hat{\beta}_\tau$ is the sample analogy of Equation 3. This minimization procedure involves the solution of a linear programming problem. As one increases τ from 0 to 1, one traces the entire conditional distribution of y_{it}, conditional on x_{it}.

5.2. Wage - productivity levels

In the first model we mean to detect the relationship between the level of productivity and the level of wages. The model, estimated at the highest levels of sectoral disaggregation (four-digit) at the $0.05, 0.10, 0.25, 0.50, 0.75, 0.90, 0.95$ quantiles of the conditional wage *level* distribution, reads as:

$$w_{it} = \alpha + \beta_\tau \pi_{it} + y_t + \epsilon_{\tau it} \tag{4}$$

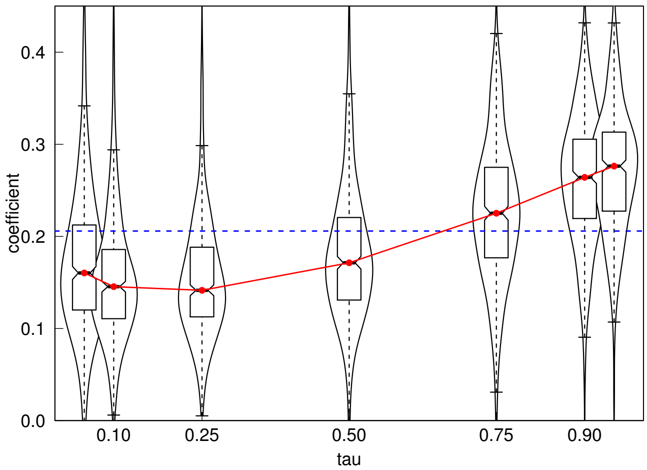
where w_{it}, the dependent variable, is the (log) real wage per employee for firm i at time t, π_{it} the (log) labour productivity level. We also control for common macroeconomic shocks by including year dummies y_t. The violin plot in Figure 15 presents the median, the interquartile ranges, and the kernel density distribution of the coefficient estimates for each quantile of the conditional wage distribution. The estimated coefficients present a monotonic increasing pattern, meaning that the wage-productivity nexus increases along the conditional wage distribution.[7]

At the lower quantiles of the conditional wage distribution the coefficients on labour productivity are the lowest and conversely they are remarkably higher at the upper tail of the conditional wage distribution, wherein 10% increase in labour productivity tends to raise the $90th$ and $95th$ quantiles of wage distribution by 2.64% and 2.76% respectively. We have performed the nonparametric Kruskal-Wallis test (Kruskal and Wallis, 1952) to detect the median differences across the distributions at the seven quantiles of the conditional wage distribution in Figure 15. Upon rejection of the null hypothesis of this test, we conduct post hoc multiple pairwise comparisons for stochastic dominance or median difference using Dunn's test (Dinno, 2015) with a Bonferroni adjustment. The medians at the $90th$ and $95th$ are significantly higher than the medians at the $5th$, $10th$, $25th$ and $50th$ quantiles.

Based on the coefficient estimates from the quantile regression at the 4-digit sectoral level, one may predict the conditional quantile functions of wage that are at $10th$ and $90th$ percentiles of the sample productivity distribution. Figure 16 presents the estimated conditional wage distribution for the 4-digit sectors characterised by the highest number of observations. Indeed one does not find any strong regularity of wage dispersion at the $10th$ and $90th$ percentiles of productivity distribution, with some sectors displaying higher dispersion at the $90th$ rather than at the $10th$ percentile and others

[7]The observed monotonic increasing pattern is robust to higher levels of aggregation with estimates of Equation (4) pooling all manufacturing firms and including 2-digit sectoral dummies. Results are available upon request.

Figure 15: Distributions of quantile regression coefficients across 424 four-digit sectors. Note: quantile regression estimation of equation (4) for each 4-digit sector, the coefficient of log- labour productivity reported for the 0.05, 0.10, 0.25, 0.50, 0.75, 0.90 and 0.95 quantiles. Each violin reports a box plot and a kernel density to each side of the box plot. The median of Pseudo R2 is 0.1426 for quantile regression. Dashed line is the median of the distribution of OLS estimates.



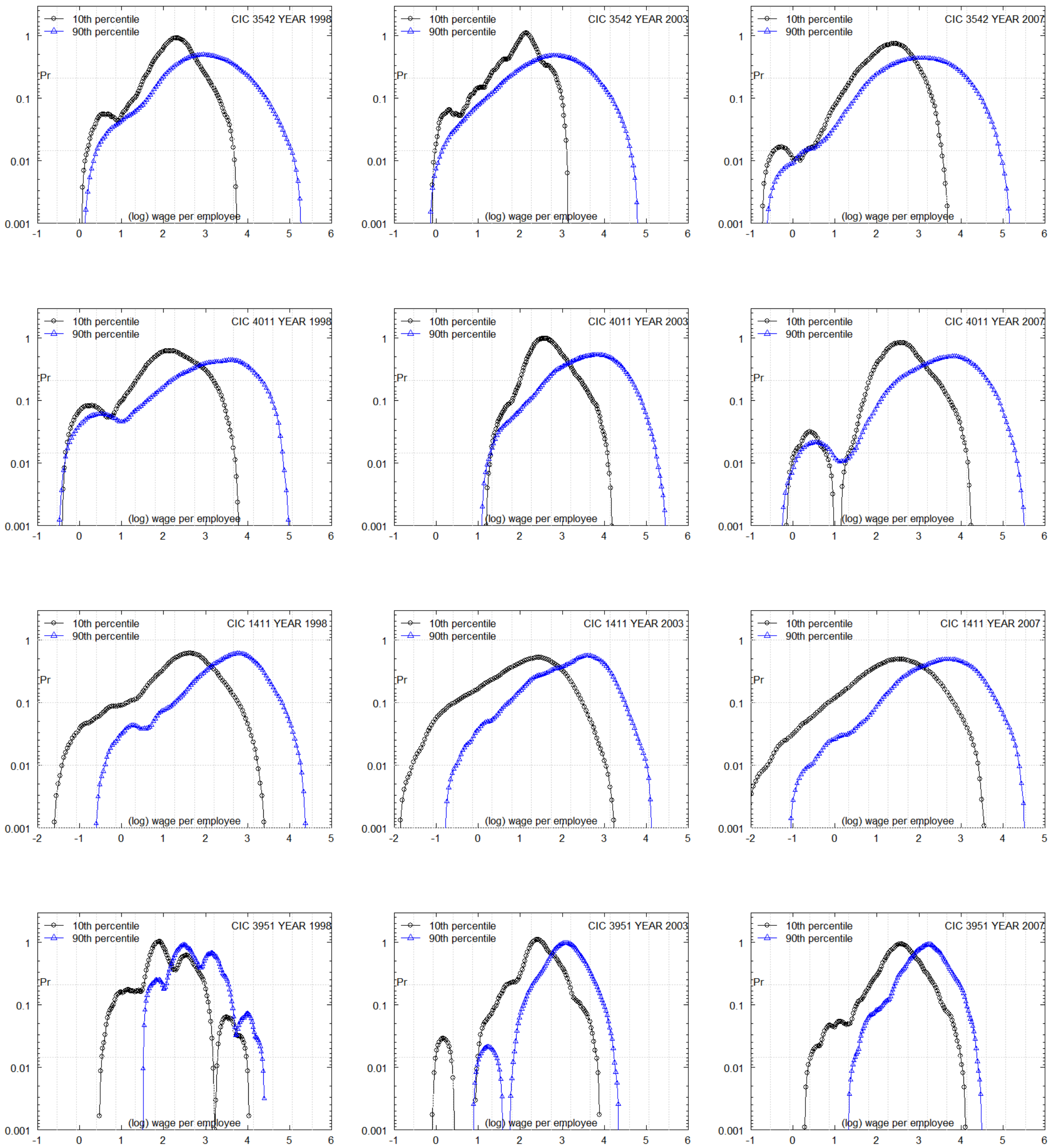
which do not. In this respect, the link between productivity levels and wage dispersion is not very robust to disaggregation by sector.

Table 7: Median of the distributions in Figure 17. Source: our elaboration on CMM.

Ownership	0.05	0.10	0.25	Quantiles 0.50	0.75	0.90	0.95	OLS	$\beta_{1,0.95} - \beta_{1,0.05}$	Pseudo R2 (Median)
State-owned	0.348	0.319	0.293	0.292	0.262	0.227	0.211	0.277	-0.137	0.1781
Collective-owned	0.173	0.134	0.116	0.119	0.154	0.203	0.229	0.162	0.056	0.1235
HMT-invested	0.137	0.138	0.161	0.194	0.221	0.251	0.267	0.201	0.130	0.1466
Foreign-invested	0.156	0.163	0.203	0.254	0.293	0.286	0.284	0.246	0.128	0.1414
Shareholding	0.162	0.144	0.147	0.169	0.212	0.249	0.262	0.198	0.100	0.1423
Private-owned	0.088	0.077	0.075	0.097	0.143	0.213	0.250	0.132	0.162	0.1316

We further extended the analysis performing six separate estimations for each major ownership type. The result of the quantile regression estimates is shown in Figure 17, while the median values of the distribution of coefficients are reported in Table 7. State-owned enterprises present the highest association between productivity and wage level with a monotonic *decreasing* relationship along the conditional quantile of the wage distribution, a pattern remarkably different from the other ownership types. In fact, the quantile regression coefficients are *highest* at the lower tail of the wage distribution and *lowest* at the upper tail. In that, SOEs seem to be comparatively more *egalitarian* in the sense that productivity appears to be passed through wages more at the lowest wage quantiles and less at the highest ones. We performed again the Kruskal-Wallis test on the median differences across the seven distributions for each ownership type in Figure 17. For SOEs, the null hypothesis are always rejected. The degree of pass-through are statistically different across different quantiles on the conditional wage distributions. The opposite applies to foreign-, HMT-invested, shareholding and private-owned firms. Based on the Kruskal-Wallis and Dunn tests, the median are not statistically different at

Figure 16: Predicted wage distribution based on the estimated conditional quantile function at the $10th$ and $90th$ percentile of labour productivity distribution, four examples, manufacturing of gas compression machinery (CIC 3542, first row), communications transmitting equipment (CIC 4011, second row), aircrafts (CIC 1411, third row) and electrical appliances (CIC 3951, forth row) year 1998, 2003 and 2007. Note: log-log scale.

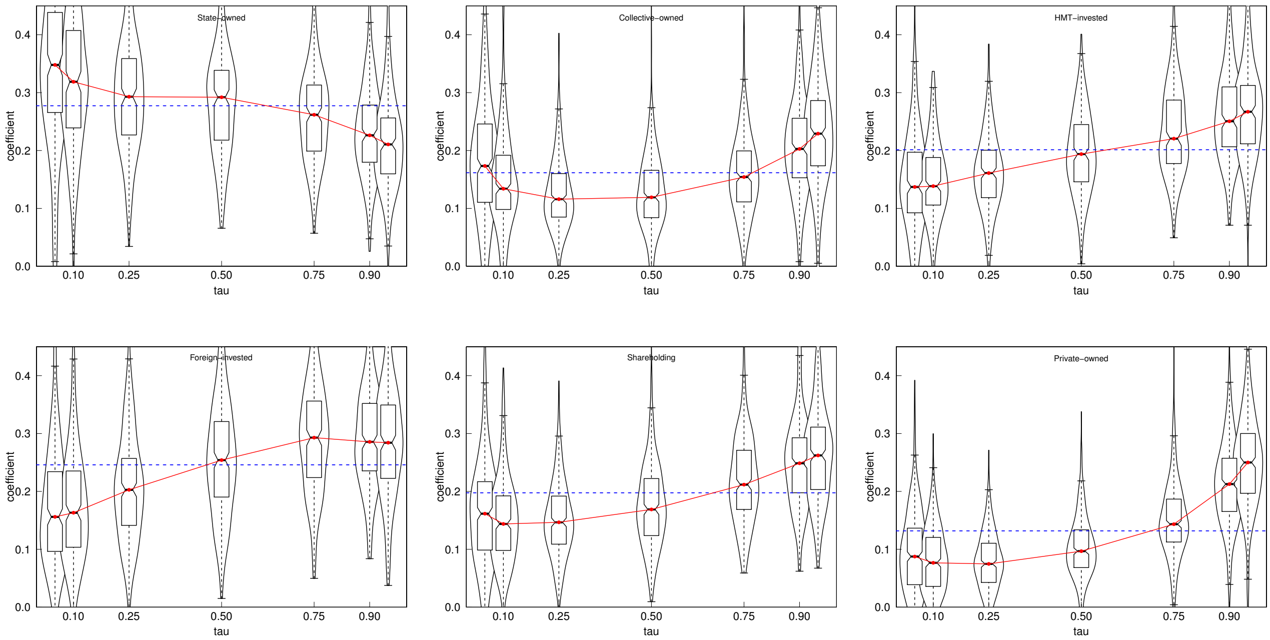


the $5th$, $10th$ and $25th$ of the conditional wage distribution. Finally, the quantile regression coefficients for collective-owned enterprises display a slightly U-shaped pattern. The link between productivity and wage is relatively higher at both the lower and the upper tails of the wage distribution, albeit there appears a much higher association at the upper tail.[8]

In order to account for the effects of other potential determinants we estimate the following model,

[8]Results are robust when estimating equation (4) for each ownership type, pooling all manufacturing firms in each ownership type, including 2-digit sectoral dummies.

Figure 17: The distribution of quantile regression coefficients across four-digit sectors for each ownership type. Quantile regression estimation of equation (4). The coefficient on log- labour productivity reported for the 0.05, 0.10, 0.25, 0.50, 0.75, 0.90 and 0.95 quantiles. Note: keep sectors with observations > 160. Dashed line is the median of the distribution of OLS estimates.



accounting for the role of size, age, export, geographical location:[9]

$$w_{it} = \alpha + \beta_{\tau 1}\pi_{it} + \beta_{\tau 2}size_{it} + \beta_{\tau 3}age_{it} + \beta_{\tau 4}export_{it} + ownership_{it} + geo_i + y_t + \epsilon_{\tau it} \tag{5}$$

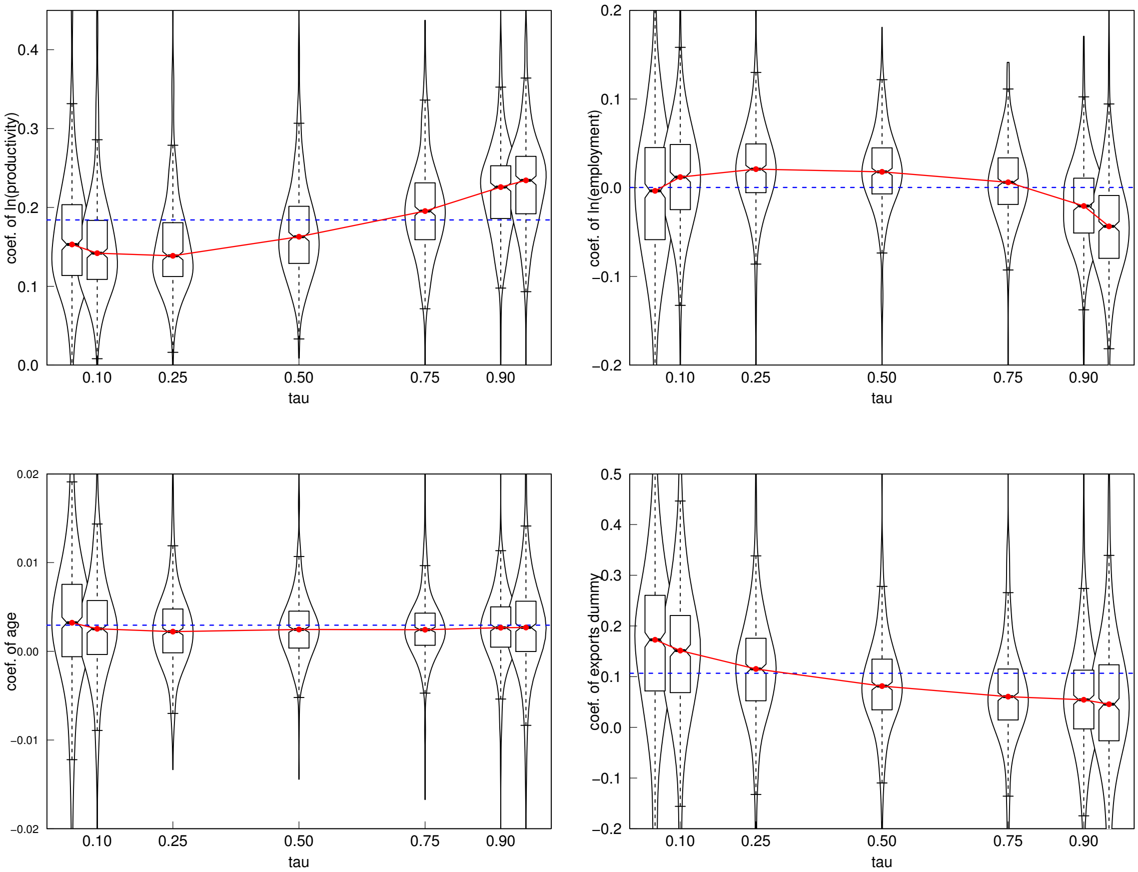
where w_{it}, the dependent variable, is the (log) real wage per employee for firm i at time t, π the (log) labour productivity level. We control for firm's contemporaneous size (i.e., proxied by log- number of employees), age, exporting status (i.e., a time varying dummy taking value one if the firm has positive exports), ownership types and regional locations.[10] We also control for common macroeconomic shocks by including year dummies y_t.

Figure 18 shows the distribution of coefficients at each quantile of the conditional wage distribution. After performing the Wilcoxon signed-ranks test (i.e., test the equality between the median of the distribution of estimated coefficients and zero), the Kruskal-Wallis and Dunn tests (the median differences from multiple distributions, and the post hoc pairwise comparisons) we find that the positive (median of the distribution significantly different from zero) and monotonic increasing wage-productivity nexus is robust even controlling for firm's size, age, exporting status, ownership types and geographic location. Conversely, the associations between size and wage captured by the median of the coefficient distribution of size are very different at different quantiles of the wage distribution: while it is not significantly different from zero at $5th$ quantile of wage distribution, it is positive and significantly different from zero at the $10th$, $25th$, $50th$, and $75th$ quantiles of the wage distribution, and negative and significant in the upper tail of the wage distribution ($90th$ and $95th$

[9]As shown in Figure A.1 the spatial distribution of firms is rather uneven across the Chinese provinces. The increasing concentration in the coastal areas is associated with the heavy flows of migrant workers.

[10]We distinguish China into four regions: east, middle, west and north-east.

Figure 18: Level model with controls, coefficients of productivity, size, age and export dummy. The distribution of quantile regression coefficients across 424 four-digit sectors. Note: quantile regression estimation of equation (5) for each 4-digit sector, the coefficient on log- labour productivity, log- number of employees, age and export dummy reported for the 0.05, 0.10, 0.25, 0.50, 0.75, 0.90 and 0.95 quantiles. Dashed line is the median of the distribution of OLS estimates.



quantiles). The association between age and wage is weakly positive and significant, slightly indicating that older firms tend to distribute higher wages, but independently from the conditional wage quantile. Moreover, there is a positive and significant association between exporting status and wage, but interestingly it is more pronounced at the lower tail of the wage distribution, and declining along quantiles.[11]

5.3. Wage growth - productivity growth

In order to account for the dynamic structure, we estimate the relationship between wage growth and productivity growth. The values of the coefficients provide an estimate of the degree of pass-through of the latter to the former. Our model specification reads:

$$\Delta w_{it} = \alpha + \beta_\tau \Delta \pi_{it} + y_t + \epsilon_{\tau it} \tag{6}$$

where Δw_{it} is the growth rate of wage per employee for firm i at time t (log difference of the wage

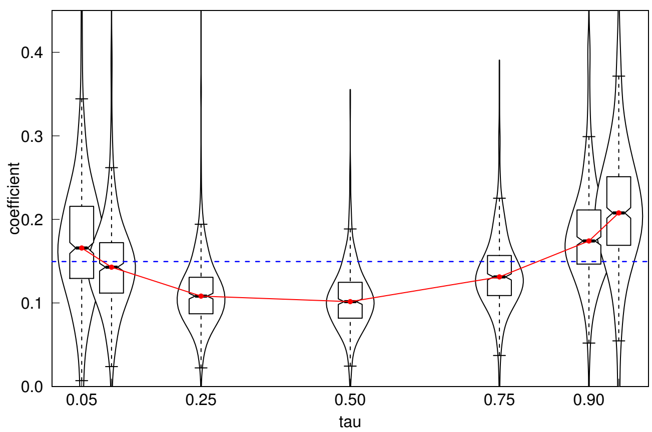
[11] We also estimate model 5 for each ownership type (excluding ownership dummies). Results are confirmed and available upon request.

at two consecutive years) and $\Delta\pi$ represents growth rate of labour productivity. We also control for common macroeconomic shocks and include year dummies y_t.[12]

We estimate model (6) for each four-digit sector at the 0.05, 0.10, 0.25, 0.50, 0.75, 0.90 and 0.95 quantiles of the conditional wage *growth rate* distribution. Results are shown in Figure 19. First notice that elasticities are all remarkably low (less than 0.21). Moreover, unlike the estimates in levels, the patterns of pass-through are U-shaped. In the median interquartile range $[0.25 - 0.75]$ of the conditional wage *growth* distribution, the coefficient on labour productivity *growth* is the lowest, 0.102, that is only one tenth of labour productivity growth is passed through. Conversely, the coefficients are significantly higher in both the lower and upper tails of the wage growth distribution.[13] The significance of the U-shaped pattern is confirmed by the Kruskal-Wallis and Dunn tests. Interestingly, the U-shaped patterns are independent from the type of ownership structure: Figure 20 shows the estimates of Equation (6) for each ownership type. Table 8 shows the median value of the coefficient distributions per each quantile. The significance of the U-shaped pattern at the ownership level is confirmed by the foregoing tests.

Finally, we studied at 4-digit sectoral level the predicted wage *growth rates* distribution at the $10th$ and $90th$ percentiles of the distribution of productivity *growth rates*. Figure 21 reports the results for two sectors, which are well representative of most of them. The Figure presents a significant overlap of the wage growth distribution: independently from being a high-growth or a low-growth productivity firm, the predicted process of wage growth is the same. The latter observation clearly militates in favour of the process of convergence presented in Section 4.

Figure 19: The distribution of quantile regression coefficients over 424 four-digit sectors - quantile regression estimation of equation (6) for each 4-digit sector: the coefficient on growth of labour productivity reported for the 0.05, 0.10, 0.25, 0.50, 0.75, 0.90 and 0.95 quantiles. Median of Pseudo R2 is 0.0348. Dashed line is the median of the distribution of OLS estimates.



We also estimate the growth model with controls and we perform, as usual, the Kruskal-Wallis and Dunn tests.

$$\Delta w_{it} = \alpha + \beta_{\tau 1}\Delta\pi_{it} + \beta_{\tau 2}size_{it} + \beta_{\tau 3}age_{it} + \beta_{\tau 4}export_{it} + ownership_{it} + geo_i + y_t + \epsilon_{\tau it} \qquad (7)$$

[12] For pooled manufacturing firms, we include 2-digit industry dummies.

[13] We observe very similar results even when pooling all manufacturing firms, including 2-digit sectoral dummies.

Figure 20: Quantile regression coefficients for each ownership type. Quantile regression estimation of equation (6) pooling all manufacturing firms, include 2-digit sectoral dummies. Dashed line is the median of the distribution of OLS estimates.

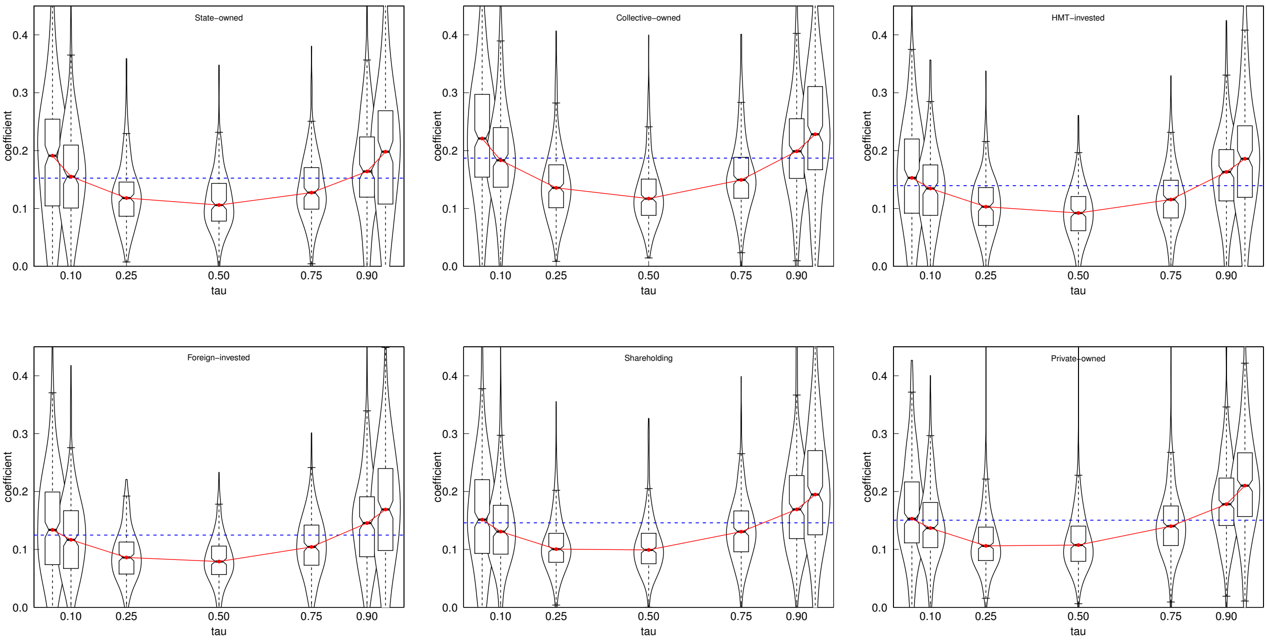


Figure 21: Predicted wage *growth rate* distribution based on the estimated conditional quantile function at the $10th$ and $90th$ percentiles of labour productivity *growth rate* distribution, two examples, manufacturing of gas compression machinery (CIC 3542, first row), communications transmitting equipment (CIC 4011, second rows), year 1998, 2003 and 2007. Note: y-axis in log scale.

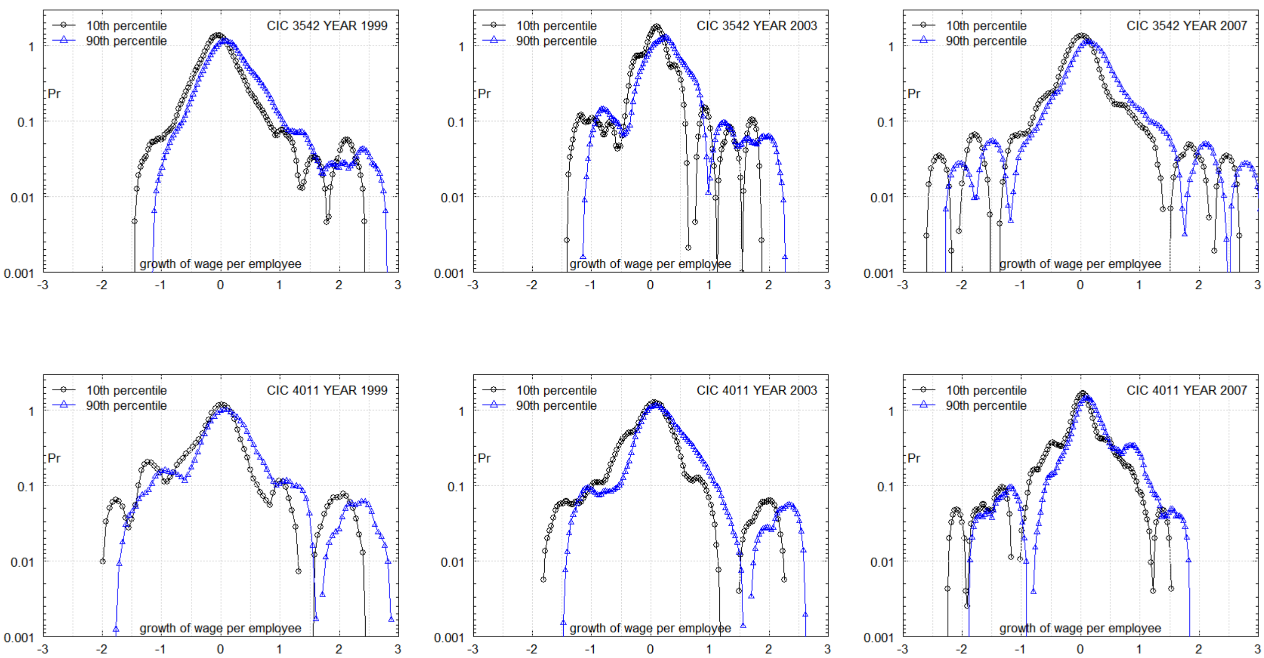


Figure 22 shows the coefficient estimates for productivity growth, size, age, and exporting status. Consistent with Figure 19, the U-shaped pattern is robust. Size displays a negative association with

Table 8: Median of the distributions in Figure 20. Source: our elaboration on CMM.

Ownership	Quantiles 0.05	0.10	0.25	0.50	0.75	0.90	0.95	OLS	Median of pseudo R2
State-owned	0.191	0.155	0.118	0.106	0.127	0.164	0.198	0.153	0.0580
Collective-owned	0.221	0.183	0.136	0.117	0.149	0.199	0.228	0.187	0.0535
HMT-invested	0.153	0.135	0.103	0.092	0.116	0.163	0.186	0.139	0.0502
Foreign-invested	0.134	0.117	0.086	0.079	0.104	0.146	0.169	0.125	0.0492
Shareholding	0.151	0.131	0.101	0.099	0.131	0.169	0.195	0.146	0.0496
Private-owned	0.153	0.137	0.106	0.108	0.140	0.178	0.210	0.150	0.0448

wage growth at the upper tail, while age has a mild decreasing monotonic pattern along the quantiles of wage growth. A similar monotonically decreasing but more pronounced pattern emerges with respect to exporting status. These aggregate findings are confirmed when decomposed by ownership structure (results available upon request).

5.4. Variance of wage and variance of productivity

In order to further study the relationship between wage and productivity dispersion, let us estimate the association between the variance of log-wage and of log-labour productivity across firms in the same 4-digit sector according to the following model:

$$\mathrm{VAR}\, w_{jt} = \alpha + \beta_\tau\, \mathrm{VAR}\, \pi_{jt} + y_t + \epsilon_{\tau jt} \tag{8}$$

where $\mathrm{VAR}\, w_{jt}$ is the variance of wage per employee (across firms) for sector j (at 4-digit level) at time t, and $\mathrm{VAR}\, \pi_{jt}$ is that of labour productivity level. The model allows to capture the extent to which wage dispersion at the sectoral level is affected by between-firms productivity dispersion. Similar OLS models have been estimated in Berlingieri et al. (2017) for a cross-country analysis. However, the quantile approach allows the study of the link between wage and productivity dispersion along the conditional distribution of between-firms wage dispersions. Figure 23 shows the quantile compared with the OLS estimates (horizontal solid line). The OLS coefficient is around 0.15. The monotonically increasing quantile pattern indicates that the higher the sectoral level wage dispersion, the higher is the contribution of sectoral productivity dispersion on it. Figure 24 shows the quantile regression and the OLS estimates (horizontal solid line) for each ownership type. The foregoing pattern is confirmed independently from the ownership structure. Note that if the generating process of wages was identical across all wage quantiles one would not observe any correlation between variances of productivity and variances in wages.

5.5. Variance of wages and mean of productivities

Yet another set of analysis aims at detecting whether the average productivity performance of the sector might affect between-firms wage dispersion, and the extent to which this might vary from less to more dispersed wage sectors. In so doing, we indirectly detect whether the wage formation mechanism is affected by some industry productivity performance. Therefore, we estimate the following model:

Figure 22: Growth model with controls, coefficients of productivity growth, size, age and export dummy. The distribution of quantile regression coefficients across 424 four-digit sectors. Note: quantile regression estimation of equation (7) for each 4-digit sector, the coefficient on log- labour productivity, log- number of employees, age and export dummy reported for the 0.05, 0.10, 0.25, 0.50, 0.75, 0.90 and 0.95 quantiles. Dashed line is the median of the distribution of OLS estimates.

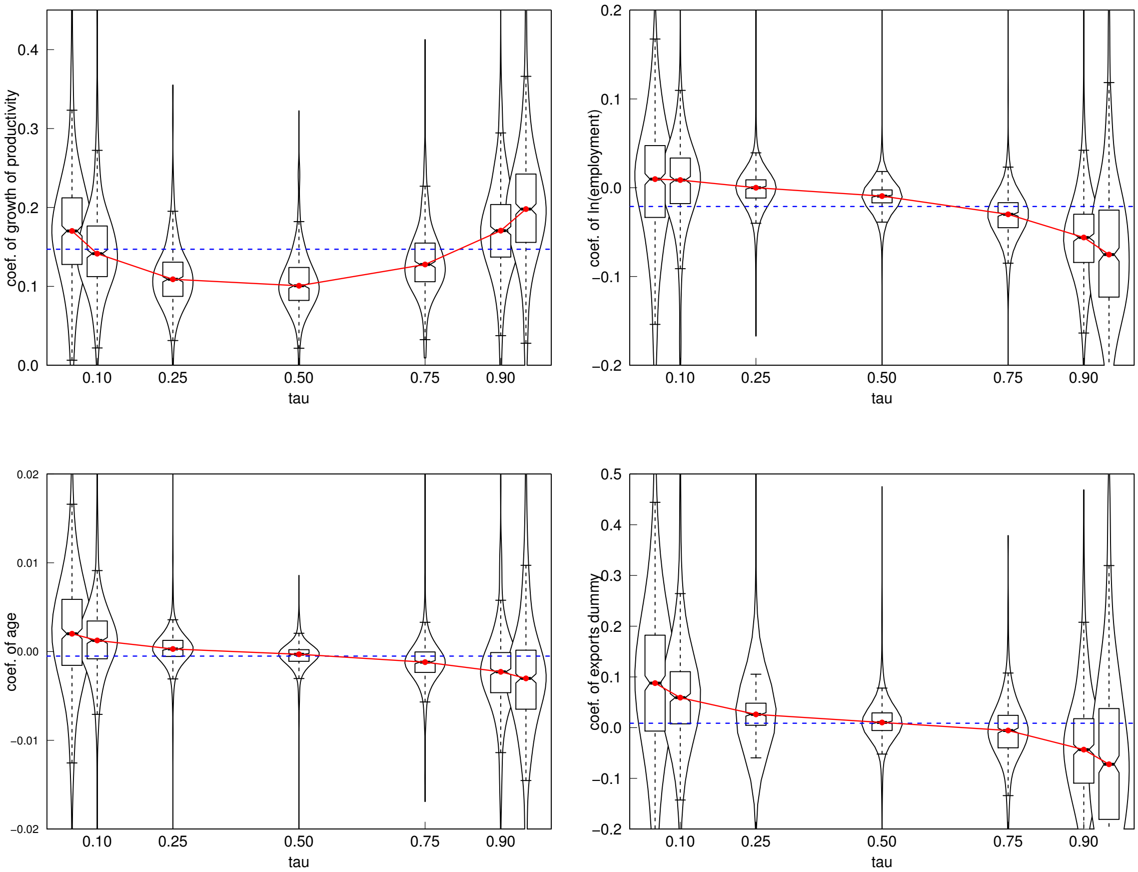


Figure 23: Quantile regression coefficients. Quantile regression estimation of equation (8) pooling all 4-digit sectors, include 2-digit sectoral dummies. Solid line is the OLS estimate.

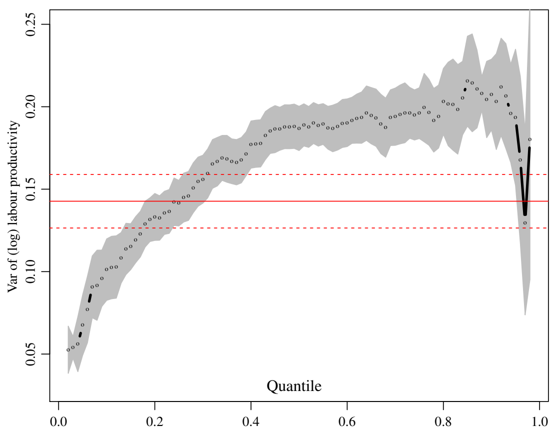
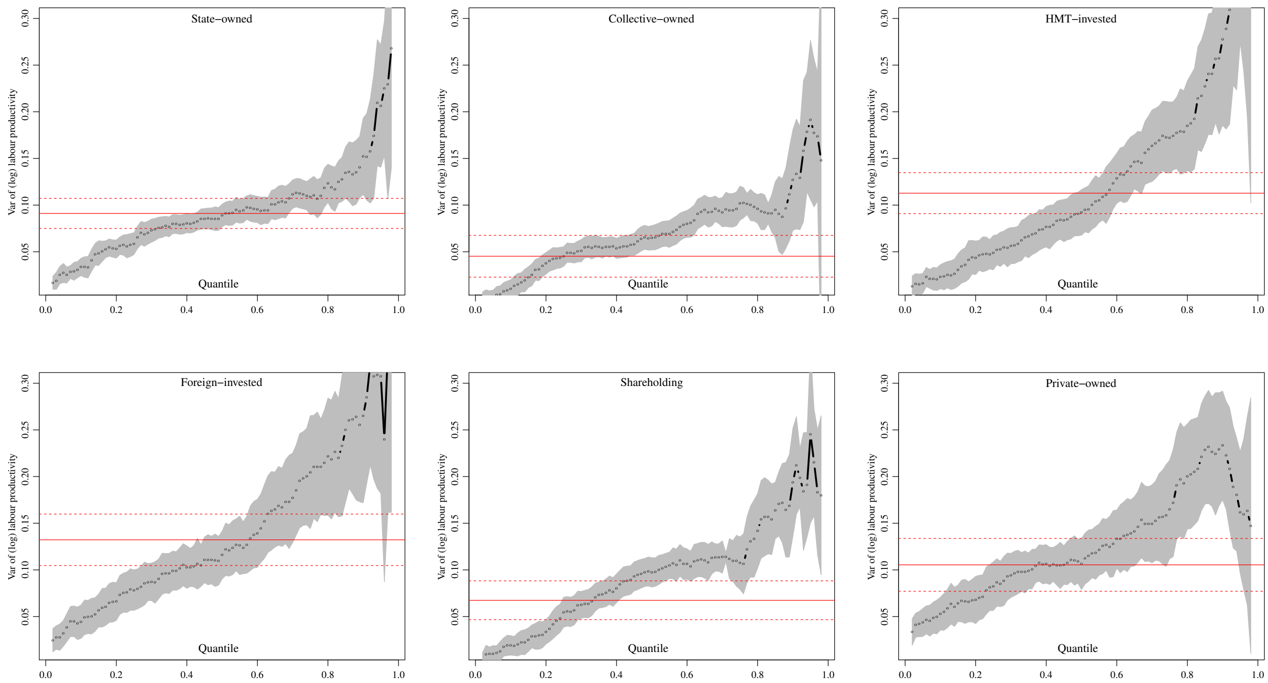


Figure 24: Quantile regression coefficients by ownership type. Quantile regression estimation of equation (8) pooling all 4-digit sectors, include 2-digit sectoral dummies. Solid line is the OLS estimate.



$$\text{VAR}\, w_{jt} = \alpha + \beta_\tau\, \text{MEAN}\, \pi_{jt} + y_t + \epsilon_{\tau jt} \tag{9}$$

where VAR w_{jt} the dependent variable, is the variance of wage per employee (across firms) for sector j (at 4-digit level) at time t. MEAN π_{jt} represents the mean of labour productivity level.

Figure 25 shows the quantile regression estimates (dashed curve) and the OLS estimate (horizontal solid line), where the OLS coefficient is around 0.075, while Figure 26 presents the same estimation for each ownership type. At the aggregate level the quantile regression analysis is redundant as the quantile coefficients and the OLS ones are largely overlapping. However when decomposing for the ownership structure, SOEs present a *negative* relationship, decreasing along quantiles, meaning that the average productivity performance of the industry exerts an equalizing effect on wage dispersion between firms, the higher the degree of wage dispersion in the given sector. The opposite occurs for the rest of the ownership types, presenting or a positive and increasing relationship along the distribution, as in the case of foreign firms (HMTs and foreign-invested enterprises), or an almost null and flat one, as in the case of domestic firms. It has to be noted that the difference with the OLS estimates looses significance for the highest quantiles in all three cases (SOEs, HMTs, foreign-invested enterprises).

5.6. Dynamic quantile regression: correlated random and fixed effects estimations

In the following we shall discuss and replicate our analysis employing the panel dimension of the data to control for unobserved heterogeneity, therefore linking quantile regression and dynamic panel techniques. In particular, we shall present the results of both the correlated random and fixed effect models. Abrevaya and Dahl (2008) propose to link the quantile regression estimation with corre-

Figure 25: Quantile regression coefficients. Quantile regression estimation of equation (9) pooling all 4-digit sectors, include 2-digit sectoral dummies. Solid line is the OLS estimate.

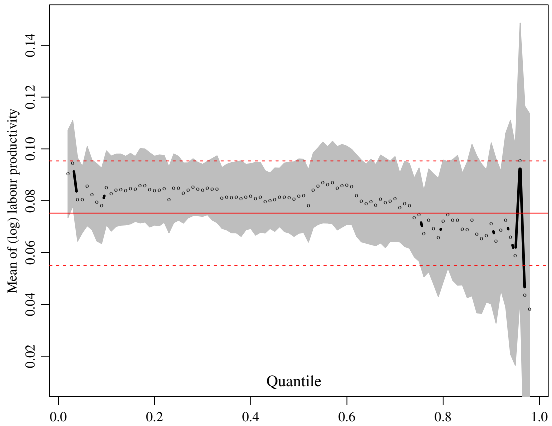
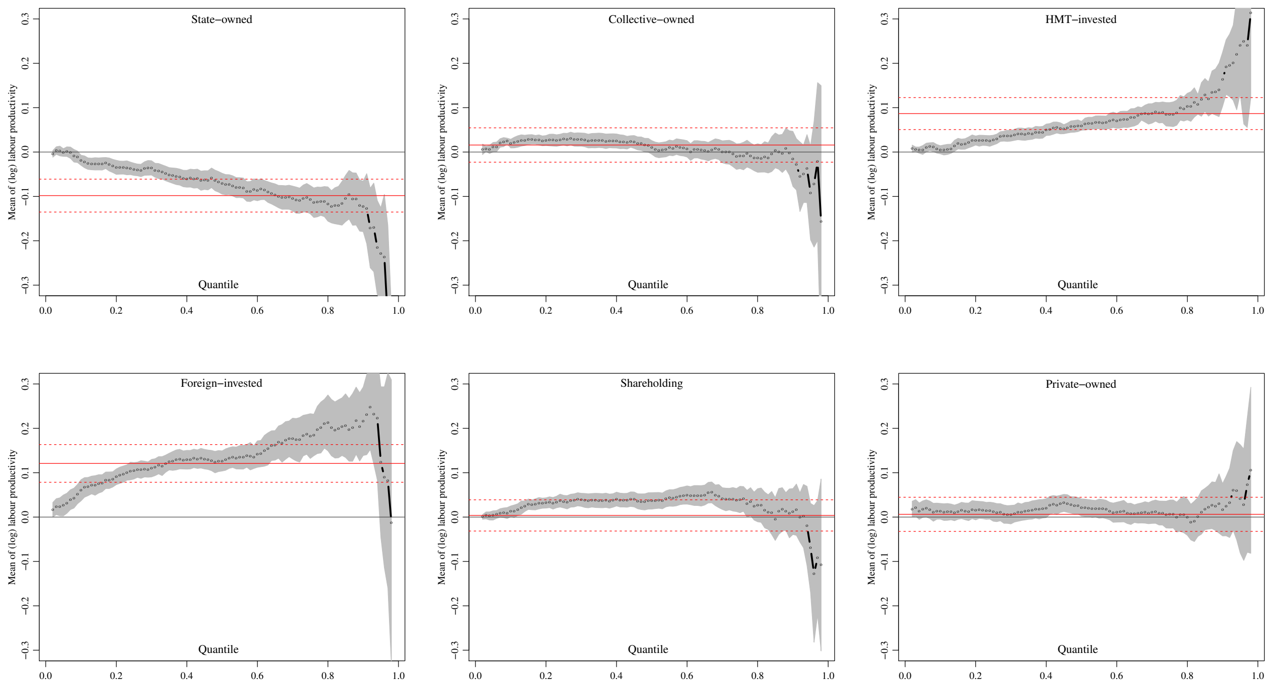


Figure 26: Quantile regression coefficients by ownership type. Quantile regression estimation of equation (9) pooling all 4-digit sectors, include 2-digit sectoral dummies. Solid line is the OLS estimate.



lated random effects using Chamberlain (1982) and Chamberlain (1984) approach (only for balanced panel).[14] According to the correlated random effect model, y_{it} is generated by:

$$y_{it} = x'_{it}\beta + c_i + u_{it} \tag{10}$$

where the time invariant idiosyncratic component c_i behaves according to:

$$c_i = \phi(x_i) + v_i, \quad E(v_i|x_i) = 0 \tag{11}$$

For any $\tau \in [0, 1]$, the conditional quantile function of y_{it} is:

$$Q_\tau(y_{it}|x_i) = x'_{it}\beta + Q_\tau(v_i + u_{it}|x_{it}) + \phi(x_i) \tag{12}$$

assuming that v_i is orthogonal to x_i and allowing for the heteroschedasticity of x_i, that is $Q_\tau(u_{it}|x_i, v_i) = Q_\tau(u_{it}|x_{it})$ we have the final specification for the quantile regression with a correlated random effect estimation:

$$\mathrm{Q}_\tau(y_{it}|x_{it}) = x'_{it}\beta_\tau + \phi(x_i) \tag{13}$$

where

$$x'_{it}\beta_\tau = x'_{it}\beta + Q_\tau(v_i + u_{it}|x_{it}) \tag{14}$$

with $\phi(x_i)$, in case of balanced panel, being:

$$\phi(x_i) = \psi^t_\tau + x'_{i1}\lambda^1_\tau + ... + x'_{iT}\lambda^T_\tau \tag{15}$$

or alternatively for unbalanced panel we have $\phi(x_i) = \psi^t_\tau + \overline{x'_i}\lambda_\tau$ (Mundlak, 1978). In the following, we estimate a wage level - productivity level quantile regression with a correlated random effect, according to such an approach, as our panel is not balanced.

Figure 27 shows the results for the correlated random effect model (CREM) which accounts for the dynamic evolution of idiosyncratic productivity over time, according to the specification of Equation 12. That is, the estimates of the wage-productivity nexus consider the micro dynamics of productivities. According to this procedure, the coefficients do not show any significant difference among quantiles at the aggregate level. Conditioning on the productivity gains the pass-through is completely flat. However, when disaggregating by ownership structure, the quantile regression with correlated random effects confirms the same pattern obtained in the pooled analysis for SOEs, as shown in Figure 30 and Table 10, with a declining pass-through across wage quantiles, but differently from the pooled estimates, an almost constant pass-through for the rest of ownership types.

An alternative specification to treat unobserved heterogeneity is the fixed effects estimator, according to Koenker (2004)'s method. The method, which proposes a penalizing estimator, imposes that the effect of the unobserved time-invariant characteristics of the firm has to be the same at each quantile τ, according to the specification:

[14] The pooled quantile regression employed the panel structure of the data only for computing standard errors. Since each firm appears at least once in the data, the clustered sampling bootstrap is used. Being present dependence within firm's indicators over years, the standard asymptotic-variance formula (Koenker and Bassett Jr, 1978) and the standard bootstrap approach, both based upon independent observations, should not be applied. Hence, instead, a given bootstrap sample is created by repeatedly drawing (with replacement) a firm from the sample of M firms and including all its measures (over years), where the draws continue until the desired bootstrap sample size is reached.

Figure 27: Model in level - CREM. The distribution of quantile regression coefficients across 424 four-digit sectors. Note: quantile regression estimation of equation (4) for each 4-digit sector, the coefficient on log- labour productivity reported for the 0.05, 0.10, 0.25, 0.50, 0.75, 0.90 and 0.95 quantiles. Dashed line is the median of the distribution of OLS estimates.

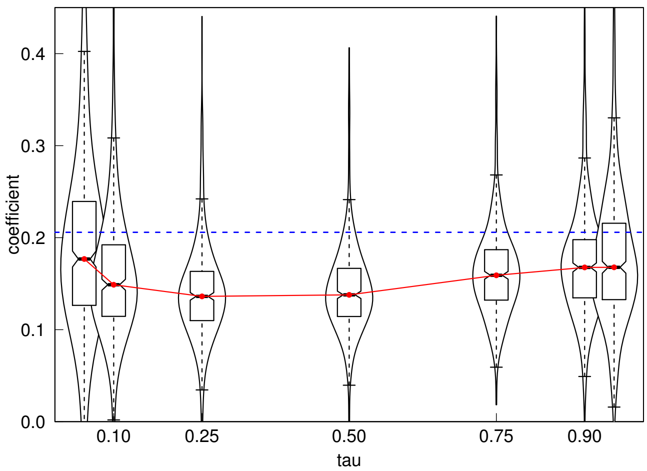


Figure 28: Model in level - CREM. The distribution of quantile regression coefficients across four-digit sectors for each ownership type. Quantile regression estimation of equation (4). The coefficient on log- labour productivity reported for the 0.05, 0.10, 0.25, 0.50, 0.75, 0.90 and 0.95 quantiles. Note: keep dataset with observations > 160. Dashed line is the median of the distribution of OLS estimates.

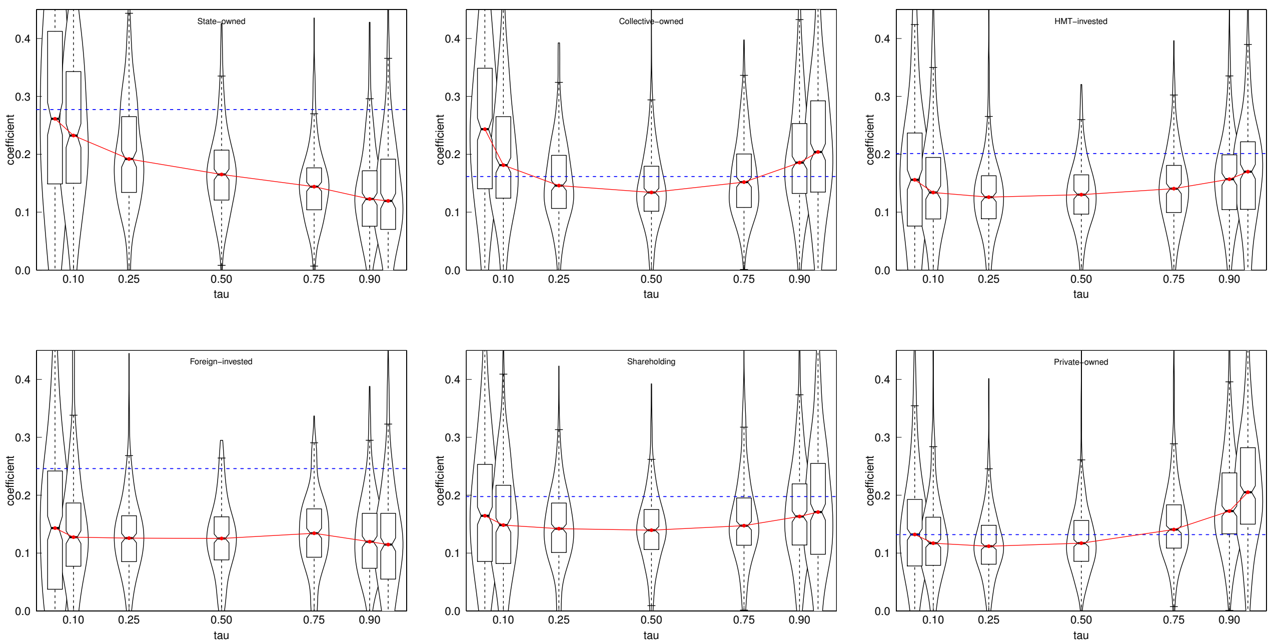
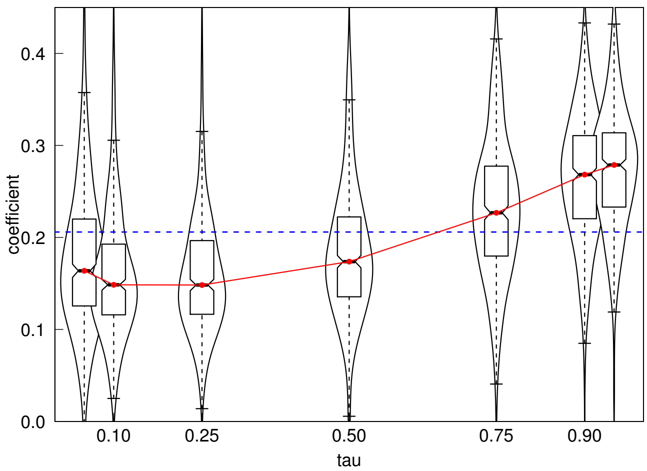


Table 9: Model in level - CREM. Median of the distributions in Figure 30. Source: our elaboration on CMM.

Ownership	0.05	0.10	0.25	Quantiles 0.50	0.75	0.90	0.95
State-owned	0.261	0.232	0.192	0.165	0.144	0.123	0.119
Collective-owned	0.243	0.181	0.146	0.134	0.152	0.186	0.204
HMT-invested	0.156	0.134	0.126	0.130	0.141	0.157	0.170
Foreign-invested	0.143	0.128	0.126	0.125	0.134	0.120	0.118
Shareholding	0.164	0.148	0.142	0.140	0.147	0.163	0.171
Private-owned	0.132	0.117	0.112	0.117	0.141	0.173	0.205

Figure 29: Model in level - FE. The distribution of quantile regression coefficients across 424 four-digit sectors. Note: quantile regression estimation of equation (4) for each 4-digit sector, the coefficient on log- labour productivity reported for the 0.05, 0.10, 0.25, 0.50, 0.75, 0.90 and 0.95 quantiles. Dashed line is the median of the distribution of OLS estimates.



$$Q_\tau(y_{it}|x_{it}) = \alpha_i + x'_{it}\beta_\tau \tag{16}$$

In this case the fixed effect is a "locational shift", not affected by the quantile. Figure 29 shows the distribution of coefficient estimates using the fixed-effects method. The FE quantile regression estimation does confirm the same pattern emerging from the pooled analysis at the aggregate level: wage dispersion increases with productivity, with an estimated magnitude of the coefficient rather similar to the pooled quantile regression. Again, similarly to the pooled analysis, when accounting for the ownership structure, SOEs show a declining pass-through along the distribution, while the opposite occurs for the rest, as shown in Figure 30. The different results provided by the CREM and the FE highlight that the pass-through of productivity gains turns out to be flat when accounting for the dynamics of firm-level productivities (CREM), washing out the apparent increasing pattern revealed by the FE estimators. In this respect, given the impressive process of productivity catching-up, the CREM reveals to be more appropriate to explicitly capture the degree of pass-through.

Finally, as a robustness test, we check the growth specification with both CREM and FE. As shown in Figure 31 in both cases the dynamic quantile specifications closely follow the model in first-differences (growth model), as expected.

Figure 30: Model in level - FE. The distribution of quantile regression coefficients across four-digit sectors for each ownership type. Quantile regression estimation of equation (4). The coefficient on log- labour productivity reported for the 0.05, 0.10, 0.25, 0.50, 0.75, 0.90 and 0.95 quantiles. Note: keep dataset with observations > 160. Dashed line is the median of the distribution of OLS estimates.

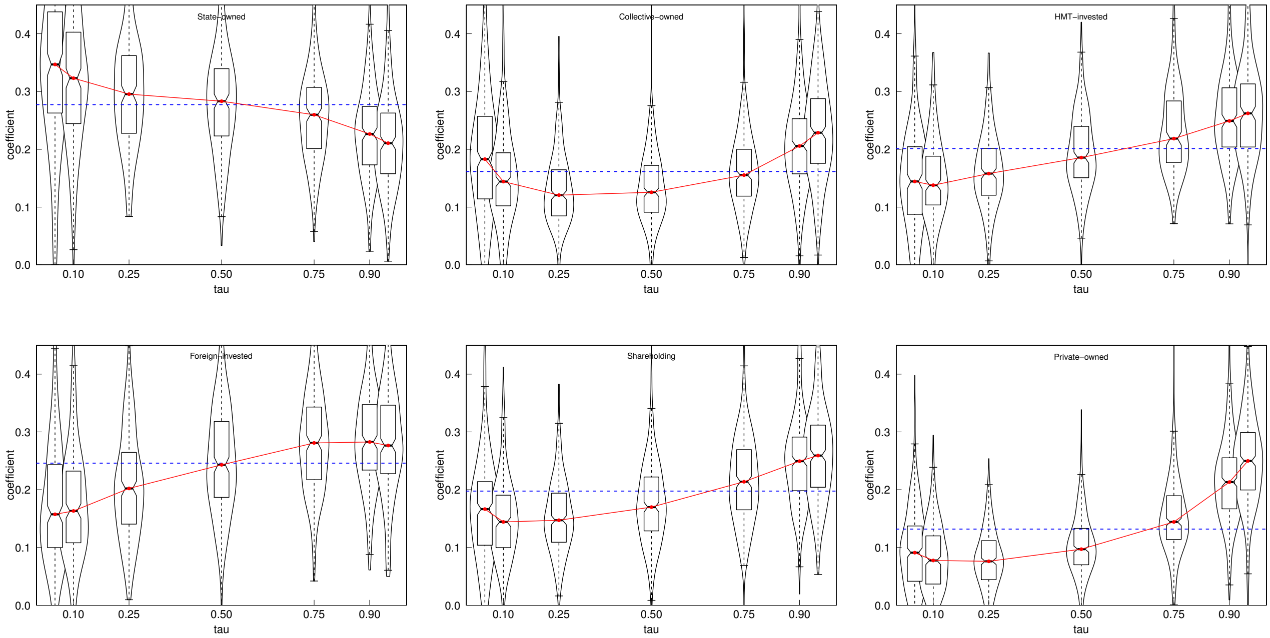
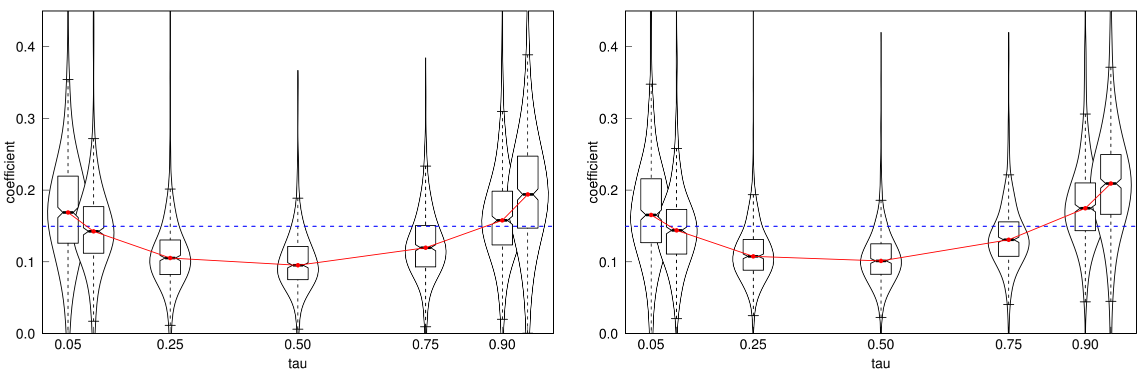


Table 10: Model in level - FE. Median of the distributions in Figure 30. Source: our elaboration on CMM.

	Quantiles						
Ownership	0.05	0.10	0.25	0.50	0.75	0.90	0.95
State-owned	0.347	0.323	0.296	0.283	0.260	0.227	0.211
Collective-owned	0.183	0.144	0.121	0.126	0.155	0.206	0.229
HMT-invested	0.144	0.138	0.158	0.186	0.219	0.249	0.262
Foreign-invested	0.157	0.163	0.202	0.243	0.281	0.283	0.277
Shareholding	0.166	0.144	0.147	0.170	0.214	0.249	0.259
Private-owned	0.091	0.078	0.076	0.097	0.144	0.213	0.250

Figure 31: Model in growth - CREM (left), FE (right). The distribution of quantile regression coefficients across 424 four-digit sectors. Note: quantile regression estimation of equation (6) for each 4-digit sector, the coefficient on growth of labour productivity reported for the 0.05, 0.10, 0.25, 0.50, 0.75, 0.90 and 0.95 quantiles. Dashed line is the median of the distribution of OLS estimates.



6. Theoretical interpretation and conclusions

Using a detailed firm-level dataset which includes all industrial firms above some minimum threshold over the period 1998-2007, distinguishing also the ownership types of the firms (e.g., State-owned, foreign MNCs, private-owned, etc.), we investigate the inter-sectoral and inter-institutional heterogeneity in both wage and productivity distributions and the coupled dynamics of the two variables, both in levels and rates of growth. Our results show a process of convergence both in productivity and wage distributions driven by a declining 50-10 percentile ratio. When decomposing the variance in terms of within and between sectoral variations, the within sectoral component accounts for more than 80% of the total variation. However, while the within sectoral dispersion in productivity shrinks, the wage one remains almost stable. We then perform quantile regressions, trying to control for different wage-productivity relations over the quantiles of the distributions. And we further refine the analysis with correlated random effects and fixed effects quantile estimations to explicitly account for the panel structure of our dataset.

A few results are quite robust under different estimation strategies. *First*, throughout the impressive Chinese economic miracle, elasticities of real wages to productivities – that is the ratios of rates of variations of the former to the latter – are always *positive* both under pooled and longitudinal estimates, both at firm- and sectoral-levels. *Second*, such elasticities are remarkably *low*, and decreasing in many sub-periods since the late seventies. The foregoing stylised facts, taken together, suggest that China has never experienced a pure *Lewis-Marx stage* of early industrialization whereby an "unlimited supply of labour" has kept wages at some subsistence level with labour productivity exponentially growing. However, even in the manufacturing sector, the distribution of gains from the impressive labour productivity growth appears to be markedly uneven. Recall that our evidence suggests that, at best, a 1% increase in productivity translates into 0.3% increase in real wages. Finally, *third*, governance institutions seem to matter a lot. So, most ownership types display (very low) firm-specific, positive elasticities of real wages to productivities. Conversely, State-owned enterprises show higher elasticities to *average* productivity growth, but basically no dependence on their own specific dynamics.

Overall, the results presented above militate in favour of the presence of two co-existing regimes of wage formation. A first one characterizing State Owned enterprises hints at the fact that firm-level wages ($w_{i_{SO}}$) are (partly) indexed on the average productivity level of the industry ($\bar{\pi}_{SO}$) rather than on the firm level one: Equation 17 captures the idea, illustrated in Figure 32.A. This implies that the pass-through is declining along the inverse of the rank of productivity distributions (cf. Equation 18 and Figure 32.B): the lower the productivity level, the higher the pass-through (γ), and the other way round.

Conversely, the wage-setting process occurring in all the rest of the firms seems rather different: it entails that wages ($w_{i_{PO}}$) are (quite partially) indexed on firm level-productivity ($\pi_{i_{PO}}$). Therefore wages increase in some proportion to firm-level productivity increases (cf. Equation 19 and see Figure 32.C) resulting into a constant pass-through along the entire productivity range of the firms (Equation 20 and Figure 32.D), independently from firms absolute ranking in terms of productivity.

$$w_{i_{SO}} = f(\bar{\pi}_{SO}) \Rightarrow \Delta w_{i_{SO}} = g(\Delta \bar{\pi}_{SO}) \tag{17}$$

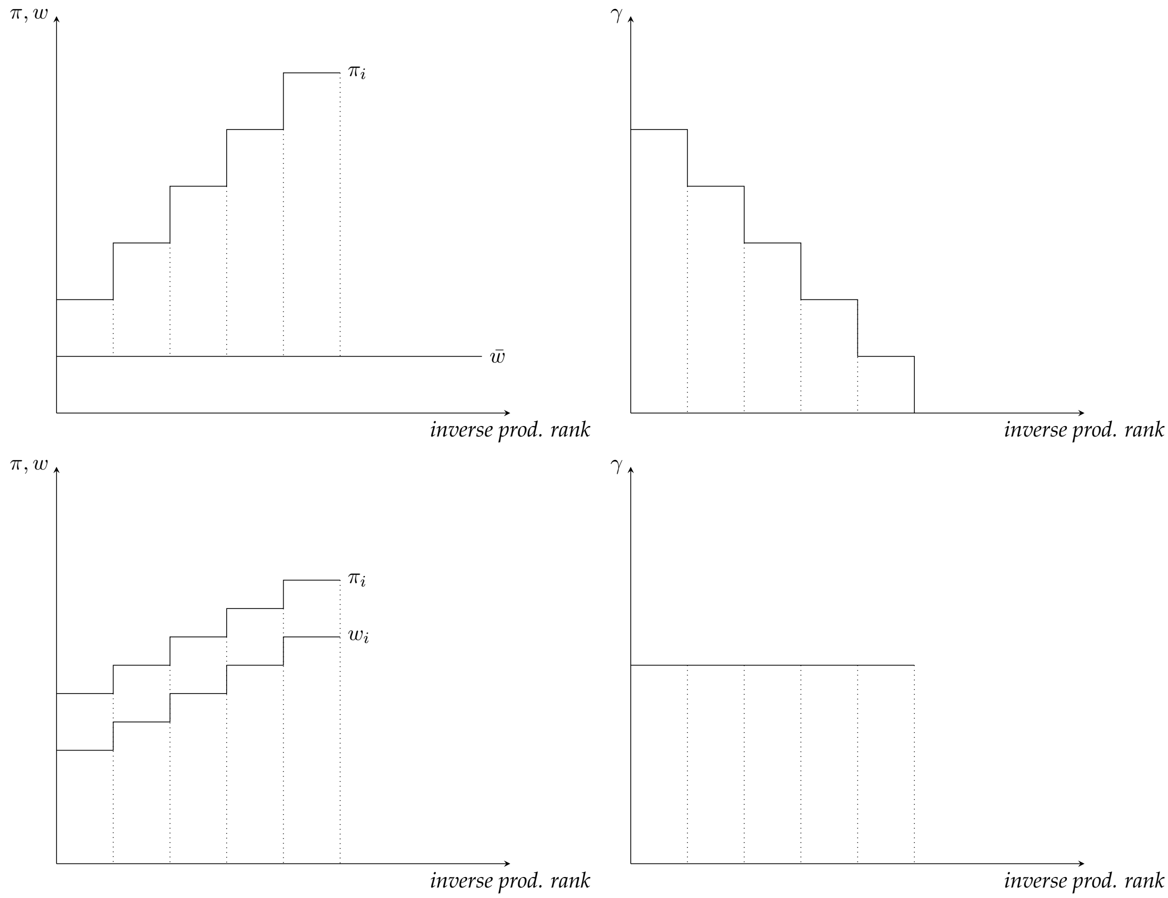


Figure 32: Top panel (State-Owned enterprises): A. The wage-productivity nexus (left); B. The ensuing pass-through (right).
Bottom panel (Private-Owned enterprises): C. The wage-productivity nexus (left); D. The ensuing pass-through (right)

$$\frac{\Delta w_{i_{SO}}}{\Delta \pi_{i_{SO}}} = h(\textit{inverse } \overset{-}{\textit{prod}}\textit{. rank}) \tag{18}$$

$$w_{i_{PO}} = f(\pi_{i_{PO}}) \Rightarrow \Delta w_{i_{PO}} = g(\Delta \pi_{i_{PO}}) \tag{19}$$

$$\frac{\Delta w_{i_{PO}}}{\Delta \pi_{i_{PO}}} = h(\textit{inverse } \overset{\perp}{\textit{prod}}\textit{. rank}) \tag{20}$$

All that, in turns, entails an underlying equalizing effect on the wage/productivity nexus. In a way, it seems that SOEs bear still some "fossil traits" of the older Central Planning period, and also the strongest resemblance to the regime of wage determination characterizing the "Golden Age" of Western post-WWII capitalist growth. Conversely, the other governance forms appear to be much more market-driven. The low elasticities of wages to productivity are plausibly the consequence of the massive flow of migrant workers from the rural areas to the coasts, somewhat resembling the early phase of the English Industrial Revolution with the pattern of enclosure in the country-side and massive migrations to the industrial towns.

Acknowledgements

The authors thank Jiasu Lei and several participants to the 30th EAEPE Conference (Nice, 2018); 17th ISS Conference (Seoul, 2018); the ICOAE Conference (Warsaw, 2018). Helpful comments have been received from William Maloney, Angelo Secchi, Paolo Piacentini, Marco Vivarelli, Fabio Montobbio, as well as seminar participants at Universita' Cattolica del Sacro Cuore (Milan, 2018), IRPET (Florence, 2018) and University of Pavia (Pavia, 2019). We gratefully acknowledge the support by the European Union Horizon 2020 Research and Innovation programme under grant agreement No. 822781 - GROWINPRO. Xiaodan Yu gratefully acknowledges the support by Martin Lockett, Martin Liu, Sathyaprakash Balaji Makam, the Nottingham Business School and the ORSS grant. The usual disclaimer applies.

References

Abrevaya, J. and C. M. Dahl (2008). The effects of birth inputs on birthweight: evidence from quantile estimation on panel data. *Journal of Business & Economic Statistics 26*(4), 379–397.

Appleton, S., L. Song, and Q. Xia (2014). Understanding urban wage inequality in China 1988-2008: Evidence from quantile analysis. *World Development 62*, 1 – 13.

Barth, E., A. Bryson, J. C. Davis, and R. Freeman (2016). It's where you work: Increases in the dispersion of earnings across establishments and individuals in the United States. *Journal of Labor Economics 34*(S2), S67–S97.

Berlingieri, G., P. Blanchenay, and C. Criscuolo (2017). The great divergence (s). Technical report, OECD Science, Technology and Industry Policy Papers, No. 39, OECD Publishing, Paris.

Brandt, L., J. Van Biesebroeck, and Y. Zhang (2012). Creative accounting or creative destruction? firm-level productivity growth in Chinese manufacturing. *Journal of Development Economics 97*(2), 339–351.

Card, D., A. R. Cardoso, J. Heining, and P. Kline (2018). Firms and labor market inequality: Evidence and some theory. *Journal of Labor Economics 36*(S1), S13–S70.

Chamberlain, G. (1982). Multivariate regression models for panel data. *Journal of econometrics 18*(1), 5–46.

Chamberlain, G. (1984). Panel data. *Handbook of econometrics 2*, 1247–1318.

Dinno, A. (2015). Nonparametric pairwise multiple comparisons in independent groups using Dunn's test. *Stata Journal 15*, 92–300.

Dong, X.-Y. (2005). Wage inequality and between-firm wage dispersion in the 1990s: A comparison of rural and urban enterprises in China. *Journal of Comparative Economics 33*(4), 664 – 687.

Dunne, T., L. Foster, J. Haltiwanger, and K. R. Troske (2004). Wage and productivity dispersion in united states manufacturing: The role of computer investment. *Journal of Labor Economics 22*(2), 397–429.

Fang, C. and D. Yang (2011). Wage increases, wage convergence, and the Lewis turning point in China. *China economic review 22*(4), 601–610.

Han, J., R. Liu, and J. Zhang (2012). Globalization and wage inequality: Evidence from urban China. *Journal of International Economics 87*(2), 288 – 297.

Hau, H., Y. Huang, and G. Wang (2016). Firm response to competitive shocks: Evidence from China's minimum wage policy. Research Paper 16-47, Swiss Finance Institute.

Jefferson, G., A. Hu, X. Guan, and X. Yu (2003). Ownership, performance, and innovation in China's large and medium- size industrial enterprise sector. *China Economic Review, Elsevier 14*(1), 89–113.

Koenker, R. (2004). Quantile regression for longitudinal data. *Journal of Multivariate Analysis 91*(1), 74–89.

Koenker, R. and G. Bassett Jr (1978). Regression quantiles. *Econometrica 46*(1), 33–50.

Koenker, R. and K. F. Hallock (2001, December). Quantile regression. *Journal of Economic Perspectives 15*(4), 143–156.

Kruskal, W. H. and W. A. Wallis (1952). Use of ranks in one-criterion variance analysis. *Journal of the American statistical Association 47*(260), 583–621.

Mundlak, Y. (1978). On the pooling of time series and cross section data. *Econometrica 46*(1), 69–85.

Schwellnus, C., A. Kappeler, and P.-A. Pionnier (2017). Decoupling of wages from productivity. Technical report, OECD Economics Department Working Papers, No. 1373, OECD Publishing, Paris.

Syverson, C. (2017, May). Challenges to mismeasurement explanations for the US productivity slowdown. *Journal of Economic Perspectives 31*(2), 165–86.

Xia, Q., L. Song, S. Li, and S. Appleton (2014). The effect of the state sector on wage inequality in urban China: 1988–2007. *Journal of Chinese Economic and Business Studies 12*(1), 29–45.

Yu, X., G. Dosi, J. Lei, and A. Nuvolari (2015). Institutional change and productivity growth in China's manufacturing: the microeconomics of knowledge accumulation and "creative restructuring". *Industrial and Corporate Change 24*(3), 565–602.

A. Firms demography and geographical location

Table A.1: Aggregation of the 23 registration categories. Source: Jefferson et al. (2003), Annex I.

Code	Ownership category		Code	Registration status
1	State-owned		110	State-owned enterprises
			141	State-owned jointly operated enterprises
			151	Wholly State-owned companies
2	Collective-owned		120	Collective-owned enterprises
			130	Shareholding cooperatives
			142	Collective jointly operated enterprises
3	Hong Kong, Macao, Taiwan-invested		210	Overseas joint ventures
			220	Overseas cooperatives
			230	Overseas wholly-owned enterprises
			240	Overseas shareholding limited companies
4	Foreign-invested		310	Foreign joint ventures
		Joint ventures	320	Foreign cooperatives
			340	Foreign shareholding limited companies
		Foreign MNCs	330	Foreign wholly-owned enterprises
5	Shareholding		159	Other limited liability companies
			160	Shareholding limited companies
6	Private		171	Private wholly-owned enterprises
			172	Private cooperatives enterprises
			173	Private limited liability companies
			174	Private shareholding companies
7	Other domestic		143	State-collective jointly operated enterprises
			149	Other jointly operated enterprises
			190	Other enterprises

The (residual) seventh category is not analyzed separately.

Table A.2: Number of firms (dataset after cleaning, and exclude firms' ownership belongs to category 7 - Other domestic ownership type)

Ownership types	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
State-owned	29171	31302	25955	21127	17893	13952	10245	9537	8083	5753
Collective-owned	41271	45231	40873	35738	32074	26870	19598	19495	16749	14115
HMT-invested	12082	14153	14807	16844	17896	19662	24253	26332	27964	26469
Foreign-invested	8595	9858	10661	12059	13500	16021	24521	27053	29599	28424
Shareholding	7867	11162	14259	19689	23292	26749	36281	41555	45956	43485
Private-owned	7981	12715	18055	31633	43318	57716	95746	112722	136294	128890
Total	106967	124421	124610	137090	147973	160970	210644	236694	264645	247136

Table A.3: Total number of employees (dataset after cleaning, and exclude firms' ownership belongs to category 7 - Other domestic ownership type). Unit: millions

Ownership types	1998	1999	2000	2001	2001	2003	2004	2005	2006	2007
State-owned	18.35	17.93	15.01	12.15	10.12	8.48	5.65	6.01	5.25	4.55
Collective-owned	10.14	10.40	9.19	7.50	6.67	5.53	3.39	3.45	2.95	2.62
HMT-invested	3.84	4.40	4.66	5.12	5.64	6.74	8.04	9.20	10.11	9.93
Foreign-invested	2.65	2.98	3.33	3.78	4.27	5.33	7.77	9.20	10.50	11.08
Shareholding	4.39	5.69	7.00	8.75	9.87	10.72	12.04	13.77	14.59	14.28
Private-owned	1.26	2.07	3.01	4.92	6.67	9.17	12.75	15.64	18.25	18.17
Total	40.63	43.46	42.20	42.22	43.25	45.98	49.63	57.28	61.65	60.63

Figure A.1: Distribution of the number of firms in manufacturing across regions in China. Source: our elaboration on CMM.