M&A, Market Power, and Macroeconomic Growth

Giorgi Piriashvili

University of Notre Dame,

Notre Dame, IN 46556, USA

Click here for the latest version

November 2022

Abstract

This paper empirically examines how mergers and acquisitions (M&A) and market power accompany permanent macroeconomic productivity increases. A structural VAR shows that technology shocks generate merger waves but do not increase aggregate markups. About 35% of the forecast error variance in M&A is associated with technology shocks, suggesting that productivity growth is a substantial part of M&A. The findings suggest that restricting M&A might have unintended negative consequences for long-run productivity growth.

Keywords: Mergers and acquisitions, aggregate growth, productivity.

JEL Classification: G34, O47.

1 Introduction

This paper empirically investigates how mergers and acquisitions (M&A) and aggregate markups accompany technology shocks. While M&A can act as a reallocation mechanism facilitating the absorption and diffusion of new technologies, it can also lead to an increase in merged entities' market power. Recently, the U.S. Congress has encouraged the Federal Trade Commission (FTC) and the antitrust division of the Department of Justice (DOJ) to increase M&A scrutiny. In response, FTC and DOJ have called for a revision of federal M&A guidelines (Rose and Shapiro, 2022). Therefore, I ask the following research question. How do M&A and market power accompany permanent productivity changes driven by technology shocks?

To answer the above question, I jointly analyze a time series of productivity growth, hours worked, aggregate M&A deal values, and markups and identify a structural VAR model using long-run recursive method as in Blanchard and Quah (1989).

My findings suggest that a technology shock that permanently increases productivity level in the long run by about one percent results in a merger wave that reallocates corporate equity equal to nine percent of GDP over a ten-year horizon. However, I find no evidence that the technology shock increases aggregate markups. The variance decomposition shows that the contribution of technology shocks to forecast error variance of M&A is almost zero at the impact and gradually accumulates to 35 percent over a 15-quarter horizon. These findings suggest that stochastic technological growth is a substantial part of M&A.

My results motivate caution when changing M&A policy. In particular, stringent policies might limit the resource reallocation mechanism leading to unintended consequences for macroeconomic growth. Furthermore, my findings suggest that merger waves in response to technology shocks are not associated with increased market power.

In general, M&A can take one of two forms. A merger is vertical when the involved firms function in the same supply chain. Researchers often regard vertical mergers as beneficial because they eliminate double marginalization and transaction costs. Horizontal mergers involve firms competing within the same market. Often, policy-makers consider them less beneficial than vertical mergers.

Regardless of merger type, researchers generally agree U.S. M&A policy has become more permissive over the past 40 years than it was from 1950 to 1980 (Whinston, 2008; Shapiro, 2019). The central issue in merger analysis is balancing the efficiency-enhancing effects of inter-firm synergies and anti-competitive price effects from market power increases.

A large body of research focuses on the causes of M&A activity waves (see, for example, Harford (2022)), and Eckbo (2014) offers a recent literature review studying characteristics of merging firms. While researchers do not necessarily agree on the waves' causes, the common understanding is that different factors drive different waves. Another part of the extant literature examines the role of economic shocks triggering reallocation through M&A waves.[1] My paper contributes to this line of research by providing empirical evidence of technology shock-driven M&A waves.

Significant theoretical microeconomic work has analyzed mergers on the granular level, particularly examining horizontal mergers' effects in Cournot models (Perry and Porter, 1985; Farrell and Shapiro, 1990; Levin, 1990; McAfee and Williams, 1992). Deneckere and Davidson (1985) analyze mergers' effects in the context of Bertrand competitions. Farrell and Shapiro (1990) show that mergers necessarily increase the merged firm's markups if they do not involve synergies and only reallocate resources.

[1] For example, Mitchell and Mulherin (1996) and Eisfeldt and Rampini (2006) study the business cycle properties of capital reallocation.

The findings in the applied microeconomic literature are mixed. Blonigen and Pierce (2016) find no evidence of firm-level gains due to mergers but a significant increase in the post-merger firm's market power. Ashenfelter et al. (2013) find a significant post-merger price increase for clothes dryers specifically. Similarly, Ashenfelter and Hosken (2010) find that four out of five mergers result in increased prices. In a study of ready-mixed concrete plants, Hortaçsu and Syverson (2007) find that vertically integrated plants are more productive than their pre-merger counterparts and enhance efficiency through economies of scope. Using data from the U.S. Census of Manufacturers, Maksimovic and Phillips (2002) find that M&A-based asset transfers lead to increased plant-level productivity and efficient allocation. However, Blonigen and Pierce (2016) more recently find no evidence of such plant- or firm-level gains but a significant increase in post-merger firms' market power. By investigating the relationship between M&A, markups and aggregate productivity growth on a macroeconomic level, my paper examines how such micro-founded aspects of M&A are aggregated at the economy-wide level.

Early studies exploring M&A activity on a macroeconomic level emphasize the way the process reallocates resources between firms. For example, Mulherin and Boone (2000) study acquisition and divestiture activities from 1990 to 1999 and find evidence that acquisitions responding to macroeconomic shocks entail restructuring and resource reallocation. Jovanovic and Rousseau (2002) analyze M&A via the q theory of investment. The theory suggests that firms increase investment when their market value divided by the cost of replacing their capital (i.e., q) increases. The researchers argue that high-q firms buy low-q firms, implying that merger waves are necessarily reallocation waves. In contrast, Rhodes-Kropf and Robinson (2008) formulate a model in which mergers are driven by complementarities between assets, and firms with similar productivity levels enjoy the greatest synergies. Nocke and

Yeaple (2007) develop a general equilibrium M&A model and focus on cross-border firm combinations versus greenfield investments.

Mitchell and Mulherin (1996) emphasize economic shocks, focusing on the 1982–1989 merger wave across 51 industries. Like this paper, their emphasis is on waves driven by shocks; however, they consider economic and deregulation shocks constructed via various measures in a cross-sectional setting, while I focus on technology shocks in a dynamic setting identified using a structural VAR model.

Eisfeldt and Rampini (2006) study the link between capital reallocation waves and aggregate shocks and show that capital liquidity is cyclical. Harford (2005) argues that clustered industry-wide shocks and sufficient capital liquidity lead to merger waves. My paper contributes to this line of research by providing evidence that technology shocks contribute substantially to merger waves.

More recently, the relevant literature has considered M&A's aggregate implications through the lens of search and matching models. David (2020) suggests M&A activity significantly impacts long-run output and consumption. Dimopoulos and Sacchetto (2017) develop a general equilibrium model of M&A along the lines of Hopenhayn (1992). The researchers show that M&A positively affects productivity by stimulating entry through valuable takeover options. However, the process negatively impacts productivity because low-productivity incumbents remain in the industry.

Cavenaile et al. (2021) explore antitrust policy's implications for aggregate growth. The researchers develop a Schumpeterian step-by-step growth model featuring industry-level Cournot product market competition. While the model implies that strict M&A policies can enhance growth and welfare, the analysis concentrates exclusively on horizontal mergers at the industry level. Garfinkel and Hankins (2011) empirically show that M&A waves are characterized by a large share of vertical mergers.

Fons-Rosen et al. (2022) also study M&A's role in aggregate growth. The researchers develop a Schumpeterian model of endogenous growth featuring M&A activity in which incumbents acquire startups. The researchers' focus is on so-called killer acquisitions that induce two effects. First, they provide incentives for startups through potentially profitable sales, positively affecting aggregate growth. Second, "killing" the startups' ideas discourages incumbents from innovating, negatively affecting growth.[2] They find that a startup acquisition ban could stimulate growth by 0.03 percentage points.

Broadly, my paper contributes to the literature studying M&A's macroeconomic role. The paper's key contribution is to empirically investigate M&A and aggregate markups' dynamic responses to aggregate productivity shocks using a structural VAR model, thereby imposing as few structural and behavioral assumptions as possible. The paper also contributes to recent U.S. policy discussions of more stringent M&A regulations by shedding light on the co-movement of M&A, aggregate market power, and productivity growth.

The next section of this paper discusses the data and empirical methodology. Section 3 presents my main findings, Section 4 discusses a possible mechanism and robustness checks, and Section 5 concludes.

2 Data and Empirical Specification

I obtained M&A data from the Institute of Mergers, Acquisitions, and Alliances (IMAA), which provides quarterly aggregated M&A deals' dollar values. IMAA is a nonprofit organization that collects data on both public and private firms using sources such as Thompson Financials and Capital IQ. I apply a variant of the U.S. Census Bureau's X-13 adjustment program to seasonally filter the raw M&A series

[2] The effect is akin to Arrow's replacement effect explaining why monopolists have less incentive to innovate than firms in competitive industries.

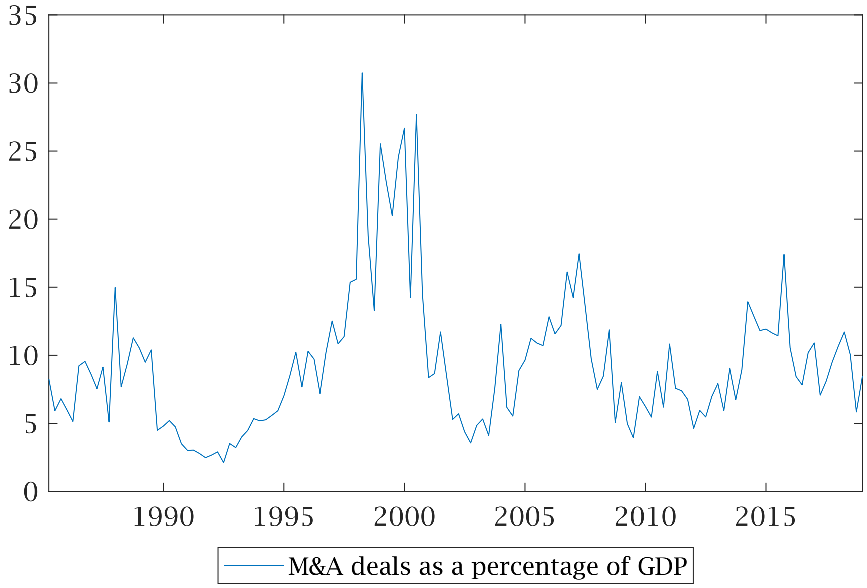


Figure 1: Time series of M&A deal values.

(Yvan, 2022).

Figure 1 plots M&A deal values as a percentage of nominal GDP for 1985-2020.[3] On average, M&A represents approximately 9% of GDP over the sample period. The peak of 30% occurs around the second quarter of 1998. The data features several well-documented merger waves.

I calculate the output growth percentages as $\Delta y_t^{NFB} = 100(\log(Y_t^{NFB}) - \log(Y_{t-1}^{NFB}))$, where Y_t^{NFB} is non-farm business output index obtained from Bureau of Economic Analysis (BEA). I measure average labor productivity as $apl_t = \log(Y_t^{NFB}) - \log(HW_t)$, where HW_t is total hours worked of all persons employed in the non-farm business sector obtained from the U.S. Bureau of Labor Statistics (BLS). The corresponding growth percentage is calculated as $\Delta apl_t = 100(apl_t - apl_{t-1})$. I measure hours worked per capita as $hw_t = \log(HW_t) - \log(POP_t)$, where POP_t is population measured in millions also obtained from the BLS.

[3] I remove 2019 and 2020 in my empirical analysis due to the COVID-19 pandemic.

I estimate the quarterly markup time series using Compustat data. I follow De Loecker et al. (2020) and use "production approach" to estimate quarterly firm-level markups. I use the weighted average of the firm-level markups, with weights determined by sales share, as an economy-wide markup measure. Appendix A reports the rest of the markup estimation details. Figure 2 plots the time series of the variables used in my analysis.

I identify a structural VAR model using the long-run recursive scheme employed by Blanchard and Quah (1989) to explore how M&A and markups co-move with technology shocks. More precisely, I use long-run recursive identification in a 4-dimensional VAR(4) model with $z_t = (\Delta apl_t, hw_t, m_t, \mu_t)'$, where Δapl_t is the percentage growth of average labor productivity, hw_t is the logarithm of total hours worked per capita, m_t is M&A as a percentage of GDP, and μ_t is de-trended aggregate markup.

To identify the technology shocks, I assume that only productivity growth shocks have a permanent effect, while the rest of the shocks are only transitory. The assumption suffices to identify structural impulse responses to productivity shocks, which is the paper's focus. I implement the exclusion restrictions by imposing a lower-triangular structure on the matrix of long-run cumulative effects.

3 Main Results

To explore how the aggregate variables co-move with technology shocks, Figure 3 (a) plots impulse responses for productivity level, hours worked, M&A, and markups to technology shocks.[4] The impulse responses are normalized to represent the effects of a technology shock that increases productivity by one percent in the long run. The

[4]Appendix B reports the full set of impulse response functions for completeness purposes.

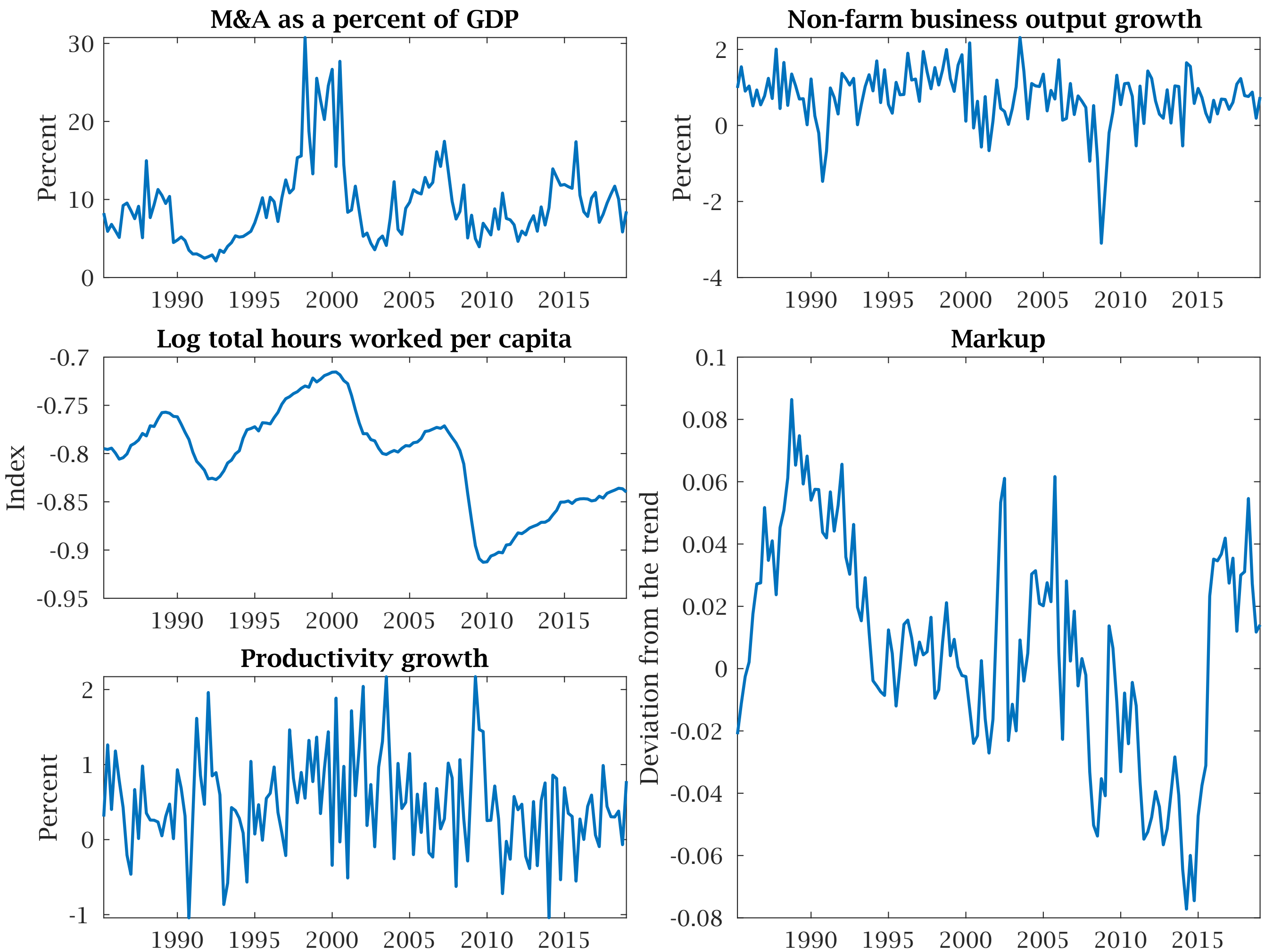


Figure 2: Time series of key variables used in the analysis.

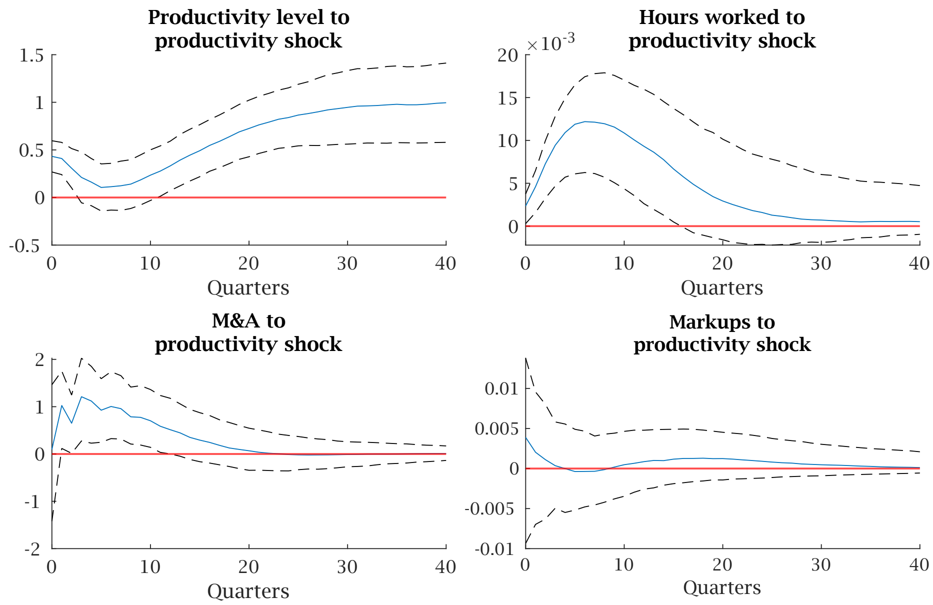
dashed lines represent 90% point-wise confidence bands.

The technology shock generates an initial increase in the productivity level, followed by a decrease for about seven quarters and gradually rising afterwards.[5] Hours worked feature a hump-shaped response with a peak at seven quarters. The M&A activity increases, reaching its peak about a year after the technology shock, and gradually declines over longer time horizons. The M&A response to the shock lags that of productivity, suggesting that M&A follows the long-run growth process, presumably through reallocating resources in the economy. Finally, the markup impulse response is not statistically significant over the 40-quarter horizon. Ignoring the confidence intervals, Figure 3 shows that the median impulse response of markup is mildly positive. The above results suggest that the merger waves generated by technology shocks are not associated with increased market power.

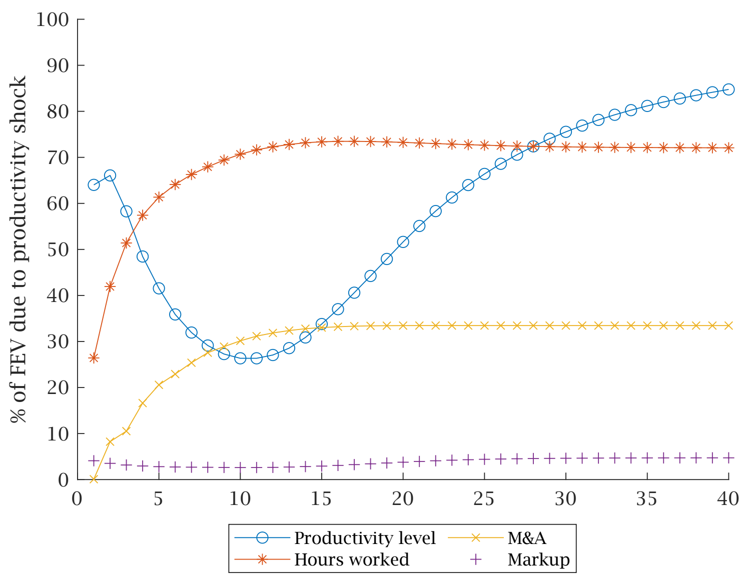
The cumulative response of M&A over the simulation horizon amounts to 9% indicating that a technology shock increasing productivity by about 1% is associated with a reallocation of corporate assets worth 9% of GDP. In my sample, labor productivity grows at about 0.5% on average; therefore, the results suggest that corporate assets worth 4.5% of annual GDP are reallocated on average in response to technology shocks.

To quantify the contribution of technology shocks in the responses of variables of interest, Figure 3 (b) reports the variance decomposition of M&A, along with total hours worked per capita, markups, and productivity. About 35% of the variation in the forecast error of M&A is due to technology shocks. Therefore, stochastic technological progress accounts for about one-third of M&A variance. Furthermore, the percentage of the variance is close to zero at the impact and gradually accumulates to

[5] Such an S-shaped pattern typically characterizes slow diffusion of technologies. See, for example, Gort and Klepper (1982).



(a) Impulse responses to technology shocks.



(b) Variance decomposition

Figure 3: Panel (a) plots impulse response functions to technology shocks increasing long-run productivity level by 1%. The dashed error bands correspond to 90% confidence intervals. Panel (b) plots the corresponding variance decomposition.

20 percent in five quarters, again suggesting that M&A follows technological progress.

Not surprisingly, technology shocks explain a sizable share of forecast error variance in productivity and hours worked. However, the shock only explains about five percent of the variance in fluctuations of markups around its trend, indicating that most markup movements are due to non-technological shocks.

To summarize the above results, technology shocks leading to permanent productivity increases generate merger waves, but I find no evidence that aggregate markups increase in response to such shocks. These findings caution when changing M&A policies. In particular, more stringent policies might negatively impact productivity growth.

4 Discussion and robustness of results

The previous section presents empirical facts regarding how M&A and markups comove with technology shocks. In this section I provide a possible mechanism and discuss the robustness of my results.

The finding that technological progress is a substantial part of M&A process suggests that long-run growth manifests itself through reallocation of resources. A possible mechanism is as follows. A technology shock generates new knowledge and ideas in the economy. The hump-shaped response of hours worked suggests that workers direct some efforts toward the absorption and implementation of new ideas. M&A activity increases to facilitate growth by reallocating resources from inefficient firms to efficient ones.

I subject my findings to a number of robustness checks. My results are robust to more conservative markup values obtained from non-parametric cost-share estimates of output elasticity. The findings are also similar to my main results when employing the utilization-adjusted total factor productivity measure from Fernald (2014). Furthermore, the results are robust to VAR order selection (VAR(2) and VAR(6) specifications yield similar results), to linear and quadratic specification of the markup trend component, and are not driven by the great recession. As the final robustness check, I investigate how the median and 90th percentile of the firm-level markup distribution respond to productivity shocks. The results are again similar to my main findings.[6]

I conclude this section by discussing two reasons for omitting the impulse responses to M&A or markup shocks. First, the key idea behind long-run identification is to view fluctuations in productivity due to two types of shocks: a shock that has permanent effects on productivity level in the long run and shocks that have only transitory effects. Intuitively, shocks that have transitory effects on productivity are identified as residuals from the shock having a permanent effect. The latter type of shock is interpreted as technology shock, while the former type – a non-technological shock.

Second, although I could identify M&A and markup shocks statistically, they would be a combination of various economic factors.[7] For example, M&A shocks could include factors like taxation-driven incentives to merge, empire-building motives, or incentives to buy startup firms to avoid future competition. Therefore, the interpretation of such shocks would be unclear.

[6] Appendix B reports impulse response functions in such specifications.

[7] This is due to the lower-triangular structure imposed on the matrix of cumulative responses discussed in Section 2.

5 Conclusion

This paper investigates how M&A and aggregate markups accompany productivity growth. I estimate a structural VAR model using long-run identification and show that technology shocks generate merger waves but do not lead to increased aggregate markups. A technology shock that permanently increases productivity by one percent results in a reallocation of corporate assets worth about 9% of GDP through M&A over a ten-year horizon. About 35% of the M&A forecast error variance is due to technology shocks indicating that stochastic technological progress accounts for a substantial part of M&A variance.

A possible mechanism, I emphasize the role of M&A in long-run growth. After a technology shock hits the economy, new knowledge or ideas become available. In response, workers direct some of their efforts to implement and absorb new knowledge. During this process, M&A acts as a resource reallocation mechanism. In particular, resources are reallocated from inefficient firms to more efficient ones.

Finally, my results caution when changing M&A-related policy. More stringent M&A policies might have unintended consequences for aggregate productivity growth. The findings are especially important in light of the recent call to increase M&A scrutiny in the United States.

References

Ashenfelter, O. and D. Hosken (2010). The effect of mergers on consumer prices: Evidence from five mergers on the enforcement margin. *The Journal of Law and Economics 53*(3), 417–466.

Ashenfelter, O. C., D. S. Hosken, and M. C. Weinberg (2013). The price effects of a large merger of manufacturers: A case study of Maytag-Whirlpool. *American Economic Journal: Economic Policy 5*(1), 239–61.

Blanchard, O. J. and D. Quah (1989). The dynamic effects of aggregate demand and supply disturbances. *The American Economic Review 79*(4), 655–673.

Blonigen, B. A. and J. R. Pierce (2016, October). Evidence for the effects of mergers on market power and efficiency. Working Paper 22750, National Bureau of Economic Research.

Cavenaile, L., M. A. Celik, and X. Tian (2021). The dynamic effects of antitrust policy on growth and welfare. *Journal of Monetary Economics 121*, 42–59.

David, J. M. (2020, 11). The Aggregate Implications of Mergers and Acquisitions. *The Review of Economic Studies 88*(4), 1796–1830.

Davis, S. J., J. Haltiwanger, R. Jarmin, J. Miranda, C. Foote, and E. Nagypal (2006). Volatility and dispersion in business growth rates: Publicly traded versus privately held firms [with comments and discussion]. *NBER macroeconomics annual 21*, 107–179.

De Loecker, J., J. Eeckhout, and G. Unger (2020, 01). The Rise of Market Power and the Macroeconomic Implications. *The Quarterly Journal of Economics 135*(2), 561–644.

De Loecker, J., P. K. Goldberg, A. K. Khandelwal, and N. Pavcnik (2016). Prices, markups, and trade reform. *Econometrica 84*(2), 445–510.

Deneckere, R. and C. Davidson (1985). Incentives to form coalitions with bertrand competition. *The RAND Journal of Economics 16*(4), 473–486.

Dimopoulos, T. and S. Sacchetto (2017). Merger activity in industry equilibrium. *Journal of Financial Economics 126*(1), 200–226.

Eckbo, B. E. (2014). Corporate takeovers and economic efficiency. *Annual Review of Financial Economics 6*(1), 51–74.

Eisfeldt, A. L. and A. A. Rampini (2006). Capital reallocation and liquidity. *Journal of Monetary Economics 53*(3), 369–399.

Farrell, J. and C. Shapiro (1990). Horizontal mergers: An equilibrium analysis. *The American Economic Review 80*(1), 107–126.

Fernald, J. G. (2014). A quarterly, utilization-adjusted series on total factor productivity. Federal Reserve Bank of San Francisco Working Paper 2012-09.

Fons-Rosen, C., P. Roldan-Blanco, and T. Schmitz (2022). The effects of startup acquisitions on innovation and economic growth. Available at SSRN: https://ssrn.com/abstract=3785485 or http://dx.doi.org/10.2139/ssrn.3785485.

Foster, L. S., J. C. Haltiwanger, and C. Tuttle (2022, September). Rising markups or changing technology? Working Paper 30491, National Bureau of Economic Research.

Garfinkel, J. A. and K. W. Hankins (2011). The role of risk management in mergers and merger waves. *Journal of Financial Economics 101*(3), 515–532.

Gort, M. and S. Klepper (1982). Time paths in the diffusion of product innovations. *The Economic Journal 92*(367), 630–653.

Harford, J. (2005). What drives merger waves? *Journal of Financial Economics 77*(3), 529–560.

Harford, J. (2022, October). Merger waves: A survey of the literature. Available at SSRN: https://ssrn.com/abstract=4244184 or http://dx.doi.org/10.2139/ssrn.4244184.

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica 60*(5), 1127–1150.

Hortaçsu, A. and C. Syverson (2007). Cementing relationships: Vertical integration, foreclosure, productivity, and prices. *Journal of Political Economy 115*(2), 250–301.

Jovanovic, B. and P. L. Rousseau (2002). The q-theory of mergers. *American Economic Review 92*(2), 198–204.

Levin, D. (1990). Horizontal mergers: The 50-percent benchmark. *The American Economic Review 80*(5), 1238–1245.

Maksimovic, V. and G. Phillips (2002). Do conglomerate firms allocate resources inefficiently across industries? Theory and Evidence. *The Journal of Finance 57*(2), 721–767.

McAfee, R. P. and M. A. Williams (1992). Horizontal mergers and antitrust policy. *The Journal of Industrial Economics 40*(2), 181–187.

Mitchell, M. L. and J. Mulherin (1996). The impact of industry shocks on takeover and restructuring activity. *Journal of Financial Economics 41*(2), 193–229.

Mulherin, J. and A. L. Boone (2000). Comparing acquisitions and divestitures. *Journal of Corporate Finance 6*(2), 117–139.

Nocke, V. and S. Yeaple (2007). Cross-border mergers and acquisitions vs. greenfield foreign direct investment: The role of firm heterogeneity. *Journal of International Economics 72*(2), 336–365.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6), 1263–1297.

Perry, M. K. and R. H. Porter (1985). Oligopoly and the incentive for horizontal merger. *The American Economic Review 75*(1), 219–227.

Rhodes-Kropf, M. and D. T. Robinson (2008). The market for mergers and the boundaries of the firm. *The Journal of Finance 63*(3), 1169–1211.

Rose, N. L. and C. Shapiro (2022). What next for the horizontal merger guidelines? Antitrust Magazine, Spring 2022 (forthcoming) URL http://dx.doi.org/10.2139/ssrn.4034923.

Shapiro, C. (2019). Protecting competition in the american economy: Merger control, tech titans, labor markets. *Journal of Economic Perspectives 33*(3), 69–93.

Whinston, M. D. (2008). *Lectures on Antitrust Economics*. The MIT Press.

Yvan, L. (2022). X-13 toolbox for seasonal filtering. https://www.mathworks.com/matlabcentral/fileexchange/49120-x-13-toolbox-for-seasonal-filtering. [Online; accessed February 21, 2022].

A Markup estimation

The markup series in the main body of the paper is constructed following De Loecker et al. (2020). I provide further details on the data and estimation procedure in this appendix.

A usual definition of the markup is output price divided by marginal cost. Empirically, it is notoriously hard to measure markups since the data on marginal cost or price is frequently unavailable. In general, there are three ways to estimate markups.

First, the so-called accounting approach assumes that the profit margins are observable. The idea is to start with the usual definition of markups, price-to-marginal cost ratio:

$$\mu \equiv \frac{P}{c} \tag{A.1}$$

and multiply throughout by quantity of output, Q, to arrive at:

$$\mu = \frac{PQ}{cQ}. \tag{A.2}$$

The simplicity of the accounting approach relies on the assumption that cQ is the total cost associated with the production and that it is directly observable in the data. The main simplifying assumption here is that marginal cost equals average cost since only then cQ is the total cost given that c is the marginal cost of a unit of output. The implication amounts to assuming that the production function exhibits constant returns to scale (CRS) and that there are no fixed costs.

Second, the demand approach involves estimating the demand system. The approach relies on having separate data on prices and quantities, an assumption on market boundaries of competing firms and a stand on the mode of competition among firms. The markups are estimated by backing out marginal cost, c, from the first-order necessary conditions of firms' profit maximization problem.

The third approach is the so-called production approach. It does not rely on assumptions on the demand side and conduct or that firms' production technology exhibits constant returns to scale. Instead, it aims to estimate the price-to-marginal cost ratio by estimating the production function directly from the first-order condition of the cost minimization problem.

Suppose there are N firms in an economy, indexed by $i \in 1, \ldots, N$. There is heterogeneity across firms and over time in terms of their productivity level Ω_{it} and possibly in terms of production technology described by production function $Q_{it}(\cdot)$. Every period each firm i minimizes the cost of production associated with producing $\bar{Q}_{it}$ units of good by choosing a vector of variable inputs, $\boldsymbol{V}_{it}$.[8] The cost minimization problem is given by:

$$\min_{\boldsymbol{V}_{it}} \sum_{j=1}^{J} P_{it}^{j} V_{it}^{j} + r_{it} K_{it} + F_{it} \qquad s.t. \qquad Q_{it}(\Omega_{it}, \boldsymbol{V}_{it}, K_{it}) = \bar{Q}_{it}, \tag{A.3}$$

where V_{it}^{j} is quantity of variable input j and P_{it}^{j} is its price and F_{it} represents fixed costs. r_{it} is the user cost of capital and K_{it} is the stock of capital that is a state variable (as opposed to a choice variable). The key assumption is that while the vector of variable inputs, $\boldsymbol{V}_{it}$, is free to adjust within a period (a quarter in my application), the capital is subject to adjustment costs or other frictions.[9] The Lagrangian associated with the cost minimization problem is:

$$\mathcal{L}(\boldsymbol{V}_{it}, K_{it}, \lambda_{it}) = \sum_{j=1}^{J} P_{it}^{j} V_{it}^{j} + r_{it} K_{it} + F_{it} - \lambda_{it}(Q_{it}(\Omega_{it}, \boldsymbol{V}_{it}, K_{it}) - \bar{Q}_{it}), \tag{A.4}$$

[8] A typical example of components of such vector would be labor, materials, energy, intermediate inputs and so on.

[9] The idea is that capital faces adjustments costs or it takes time to build. Therefore, the firms will choose variable inputs to minimize costs, given the stock of capital chosen in the previous period.

where λ_{it} is Lagrange multiplier and hence the marginal cost of production. First-order necessary condition with respect to j variable input is given by:

$$P_{it}^{j} = \lambda_{it} \frac{\partial Q_{it}}{\partial V_{it}^{j}}. \tag{A.5}$$

Multiplying throughout by V_{it}^{j}/Q_{it} and recalling the definition of markup, $\mu_{it} \equiv P_{it}/\lambda_{it}$, where P_{it} is the price of output, results in:

$$\frac{P_{it}^{j} V_{it}^{j}}{P_{it} Q_{it}} = \frac{1}{\mu_{it}} \frac{\partial Q_{it}}{\partial V_{it}^{j}} \frac{V_{it}^{j}}{Q_{it}} \tag{A.6}$$

and finally, solving for markup yields:

$$\mu_{it} = \theta_{it}^{j} \frac{P_{it} Q_{it}}{P_{it}^{j} V_{it}^{j}} \tag{A.7}$$

where θ_{it}^{j} is elasticity of output with respect to variable inpuwhere θ_{it}^{j} is elasticity of output with respect to variable input j.[10] In most of the firm-level datasets (as in case of Compustat data) sales, $P_{it}Q_{it}$, and expenditures on variable input, $P_{it}^{j}V_{it}^{j}$ are available. Therefore, the above expression can be used to back out the markup value once an estimate elasticity term is available. Subsection A.2 provides further details on estimation procedure of the elasticity term, θ_{it}^{j}.

[10]Notice that by cost minimization first-order conditions should hold for each variable input j. It is up to the researcher to choose an input that arguably best fits the definition of the variable input. In a data-rich environment where multiple variable inputs are available, one could essentially test whether markups implied by different first-order conditions are equal.

A.1 Data

I use the Compustat dataset to estimate markups. To the best of my knowledge, the Compustat dataset is the only source that provides substantial coverage of private firms. Data collection from quarterly statements started in 1979; this allows us to obtain quarterly markup estimates for the sample period of 1985-2020. The main drawback of the dataset is that it only covers publicly traded firms and that such firms are few compared to the total number of firms. However, because publicly traded firms tend to be larger and older than non-traded firms, they account for about one-third of total employment and a much larger share of total value added in United States.[11]

A potentially serious concern regarding using such data is that it introduces selection bias in the analysis since a sample of publicly traded firms is not representative of the universe of all firms. I rely on De Loecker et al. (2020) who repeat their estimation on the universe of firms using the U.S. Census dataset and find similar to their baseline results using Compustat data.[12]

The Compustat data provides firm-level information based on the financial statements. The variable used to measure variable inputs is the cost of goods sold (COGS). The variable combines all expenses directly attributable to the production of goods the firm sells. It includes costs such as labor cost, materials, intermediate inputs, energy and so on.[13] To measure the firm-level stock of capital, I use gross capital (PPEGT) and obtain a measure of capital expenditures by multiplying the deflated value of the capital stock by the user cost of capital. I follow the standard practice in

[11] See Davis et al. (2006).

[12] The authors also use Census weights to account for any bias induced by changes in sectoral composition over time.

[13] Unfortunately, Compustat does not provide a further breakdown of the combined variables and hence does not allow one to consider production functions that explicitly include such inputs.

literature and construct user cost of capital as $r_t = i_t - \pi_t + \delta$. For nominal interest rate and inflation, i_t and π_t, I use federal funds effective rate and percentage change in GDP deflator, respectively. Finally, the depreciation rate, δ, is set to 3% quarterly.

A.2 Estimation of output elasticity

This section provides details regarding the estimation of the elasticity term θ^j_{it}. I estimate output elasticity for each two-digit NAICS industry. In particular, I assume that each industry is characterized by Cobb-Douglas production function with a variable input bundle and capital as production factors and productivity entering in Hick-neutral form. Written in log form, the production function is:

$$y_{it} = \theta^v v_{it} + \theta^k k_{it} + \omega_{it} + \varepsilon_{it}, \tag{A.8}$$

where y_{it} if logarithm of output measure of firm i at time t, k_{it} is logarithm of capital, ω_{it} is logarithm of Hicks-neutral productivity level that is observed by firm (but unobservable to econometrician) and ε_{it} is measurement error in output. Finally, θ^v is output elasticity with respect to variable input, a key object of interest for markup estimation.[14]

There are two challenges with the estimation of production functions. First, the unobserved productivity term, ω_{it}, will be correlated with variable input, resulting in biased estimates if not dealt with. Second, in most cases (as in the case of Compustat data), output prices and variable inputs are unavailable; instead, the data reports revenues and expenditures. This induces so-called omitted price variable bias. I follow De Loecker et al. (2020) to address these issues. For completeness purposes, I

[14]In the baseline specification, I estimate time-invariant industry-specific elasticity. As suggested in the production function estimation literature, I also estimate time-varying elasticity, θ^v_t, using a 5-year rolling window and find similar estimates in terms of the markups.

will provide an overview.

To address the first problem, a standard is the so-called control function approach that builds upon Olley and Pakes (1996) insights. In particular, it is assumed that the unobserved (by econometrician) productivity term, ω_{it}, can be expressed as some unknown function of firm-level observables and state variables:[15]

$$\omega_{it} = g_t(c_{it}, k_{it}, z_{it}), \tag{A.9}$$

where c_{it} is the control variable and z_{it} contains variables that generate variation in the input demand through output and input market factors. In practice, z_{it} contains the firm's market shares computed at two, three and four-digit NAICS codes.[16] Broadly, two types of controls are usually used in the estimation, static and dynamic.

In the case of static control, a nondynamic input is used as a control variable. In the context of Compustat data, COGS will play the role of such variable. The estimation procedure consists of two stages. In the first stage, the output is regressed nonparametrically on the control variable, capital and exogenous shifters:

$$y_{it} = f_t(v_{it}, k_{it}, z_{it}) + \varepsilon_{it}. \tag{A.10}$$

In the second stage it is assumed that productivity follows a Markov process, $\omega_{it} = g(\omega_{it-1}) + \xi_{it}$. The estimate of elasticity term is obtained by either running a semi-parametric non-linear regression on the productivity process:

$$\hat{f}_{it} = \theta^v v_{it} + \theta^k k_{it} + g(\hat{f}_{it-1} - \theta^v v_{it-1} - \theta^k k_{it-1}) + \xi_{it}, \tag{A.11}$$

[15] This amounts to the assumption that the unobservable term is one dimensional (a scalar) and that firm's policy function is monotone.

[16] See De Loecker et al. (2016) for details regarding this approach.

where $\hat{f}_{it}$ are fitted values obtained from the first stage, or constructing moment conditions of the form:

$$\mathbb{E}\left(\xi_{it}(\theta^v, \theta^k)\begin{bmatrix} v_{it-1} \\ k_{it} \end{bmatrix}\right) = 0, \tag{A.12}$$

where $\xi_{it}(\theta^v, \theta^k)$ is obtained by regressing $\omega_{it}(\theta^v, \theta^k)$ nonparametrically on its lag. The estimate of productivity is again obtained from the first stage as $\hat{f}_{it} - \theta^v v_{it} - \theta^k k_{it}$. The assumption identifying output elasticity with respect to variable input is that the lagged non-dynamic input does not respond to productivity shocks, but the variable input use responds to shocks contemporaneously, and that lagged and current use of variable inputs are correlated through z_{it} containing input and output market factors.

In the case of dynamic control, investment is used as a control variable. In this case, the output elasticity is estimated using the following non-linear regression:

$$y_{it} = \theta^v v_{it} + f_t(i_{it}, k_{it}, z_{it}) + \varepsilon_{it}, \tag{A.13}$$

where i_{it} is investment. The identifying assumption in this case is that the non-dynamic input, v_{it}, is chosen at time $t - b$ $(0 < b < 1)$, the investment is chosen at time t and the productivity shocks are allowed to hit firm between these two periods.

As in the case of De Loecker et al. (2020), I also find similar estimates of the output elasticity in the case of the two controls discussed above.

Next, I turn to the second issue. In particular, using revenue and expenditure data instead of physical quantities introduces unobserved input and output price variation in the error term. This correlation biases the coefficient of interest since the error term in the regression equation becomes:

$$\omega_{it} + p_{it} - \theta^v p_{it}^v - \theta^k r_{it}. \tag{A.14}$$

An approach to handle this issue is discussed in De Loecker et al. (2016). In particular, assume that the terms involving unobserved input and out prices are a function of productivity differences and demand shifters. Since, in the estimation procedure, the productivity terms are already accounted for using the control variable, this part of the variation in the error term is automatically captured. The variable z should then capture the input and output price variation. The inclusion of market shares at various levels of aggregation is meant to capture such variation and eliminate it from the error term.[17]

[17] See further details in De Loecker et al. (2016). In short, they show that under nested logit demand specification, such market shares are exact controls that capture the wedge between output price and input prices (scaled by output elasticities).

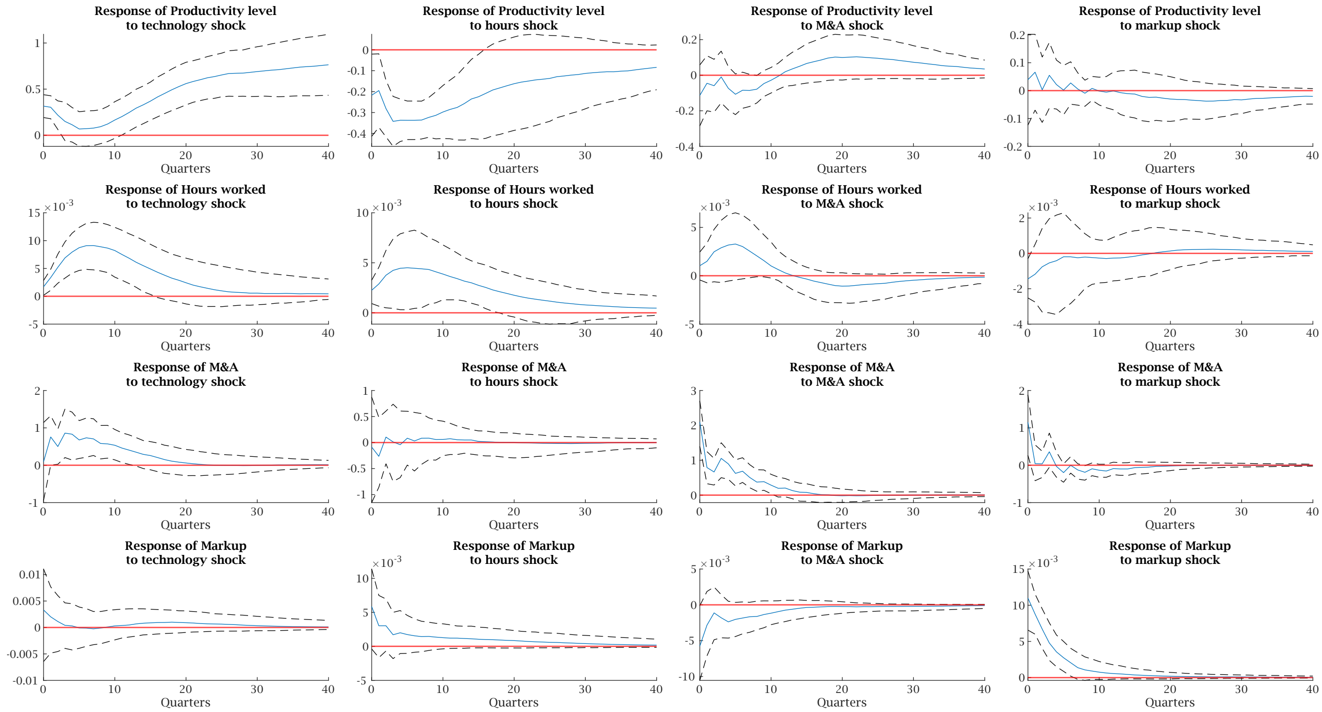


Figure 4: VAR(4) structural impulse response functions. $z_t = (\Delta apl_t, hw_t, m_t, \mu_t)'$

B Additional figures and results

This appendix contains additional figures and results omitted from the main body of the paper.

Figure 4 plots full set of impulse response functions under the main specification used in the main body of the paper. Figure 5 reports corresponding variance decompositions.

Next, I consider two additional specifications to gain further insights regarding the movements in the measure of market power. Recall that the markup estimates are at the firm level, for each time period a distribution of markups is available.[18] Therefore, one can compute various percentiles of the markup distribution for each

[18]Figure 14 plots estimated markup distributions for the first quarter of year 1990 and 2016. Notice that the right tail of the distribution increased considerably in 2016 compared to 1990.

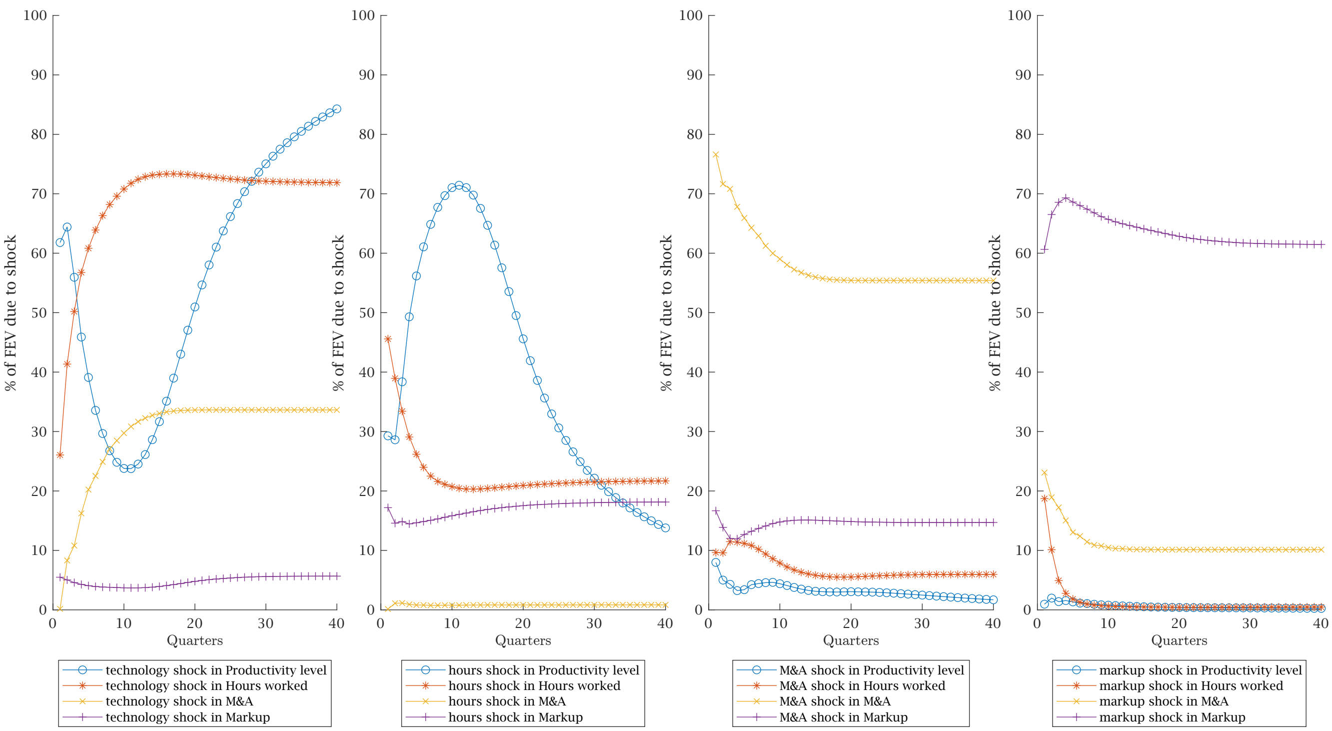


Figure 5: VAR(4) variance decomposition. $z_t = (\Delta apl_t, hw_t, m_t, \mu_t)'$

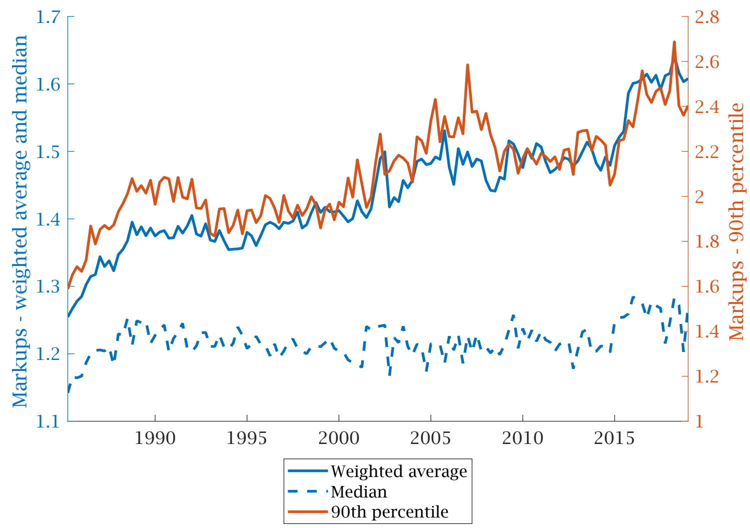


Figure 6: Quarterly markup series. The markup series are obtained by aggregating estimated firm-level markups.

time period and construct time series of markup percentiles. Figures 6 and 7 plot the time series of such percentiles and corresponding detrended series, respectively. The percentile time series are potentially useful for two reasons that I discuss below.

First, there has been concerns regarding the size of the rise in average markups as reported in De Loecker et al. (2020). For example, the rise is modest if one considers more granular level production functions.[19] To address such concerns to some extent, for each time period, I will construct median estimate of the markup distribution and use the time series constructed from the median estimates since. As suggested by Figure 6, the series of median estimates do not feature increase in markups.

Second, a plausible hypothesis could be that while the fluctuations in the average markup measure about its trend do not increase following a technology shock, the top percentiles could react differently to the increase in M&A. This is especially plausible

[19]For example, Foster et al. (2022), finds small or even no increase in markups when production functions are defined on three and four digit NAICS codes as opposed to De Loecker et al. (2020) who define production functions on two digit industry codes.

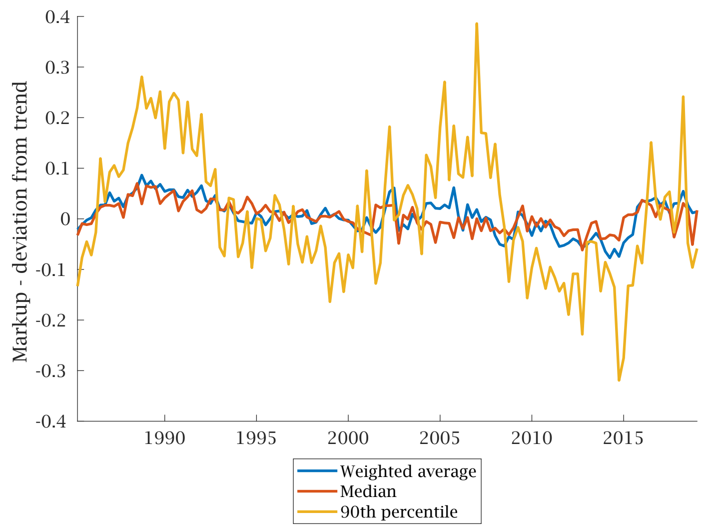


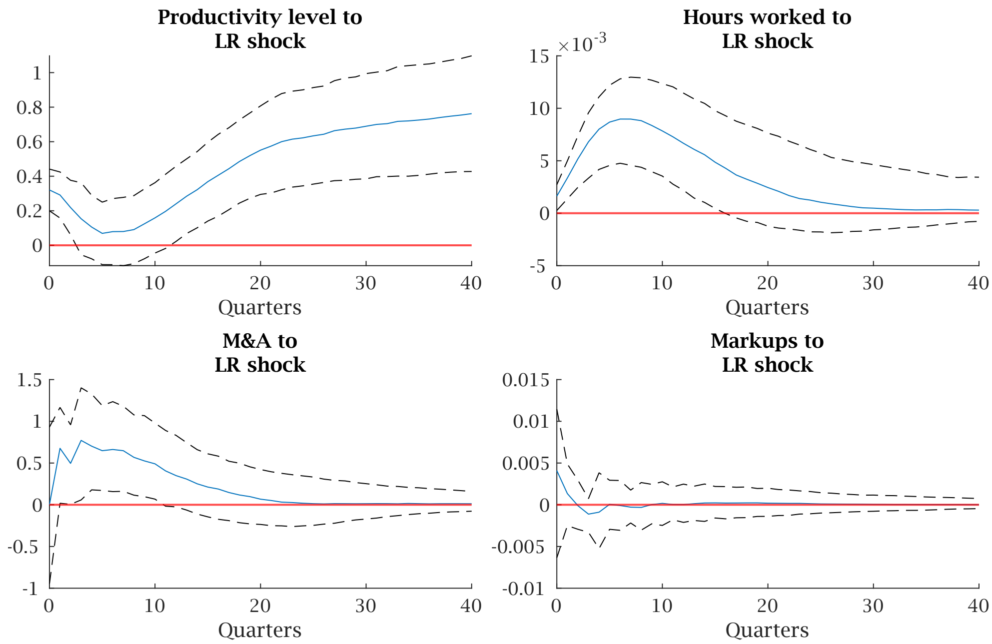
Figure 7: Quarterly detrended markup series. The markup series are obtained by aggregating estimated firm-level markups.

since, as in the case of De Loecker et al. (2020), the increase in weighted average markups is mostly driven by upper percentiles of the distribution. [20]

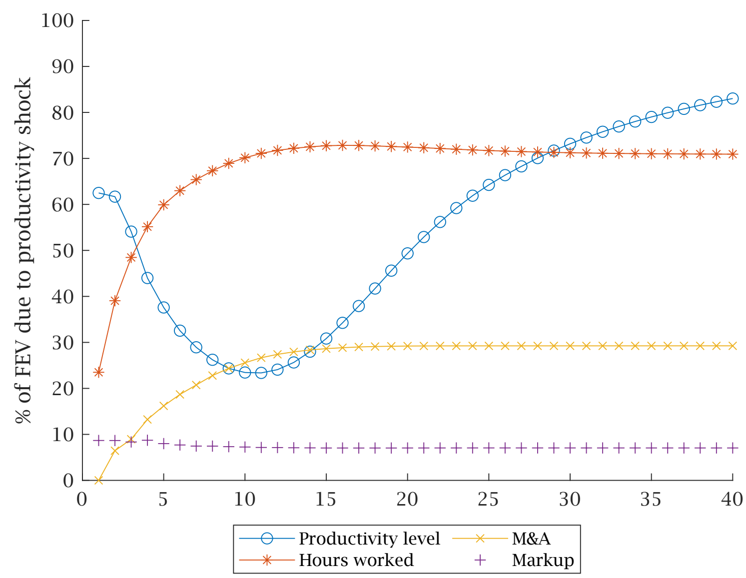
Figure 8 (a) reports impulse responses to long-run productivity shocks when detrended median markup series, μ_t^{med}, are used. We see that the pattern of impulse responses remains the same. In particular, a technology shock generates a merger wave. The response of markups exhibits a qualitatively similar pattern as in the case when the weighted average of markups was used. The bottom panel of Figure 8 plots corresponding variance decomposition. The contribution of long-run productivity shock to forecast error of M&A is still sizable. Overall, Figure 8 shows that the main findings remain the same when more conservative markup estimates (such as time series of median markups) are used.

I repeat the above exercise using the time series constructed from the 90th percentile of firm-level markup estimates to investigate how the upper part of the dis-

[20]Viewed from a different angle, the distribution of estimated markups in the late period of the sample has a larger right tail than one in the early period of the sample. See Figure 14.



(a) Impulse responses on long-run productivity shock.

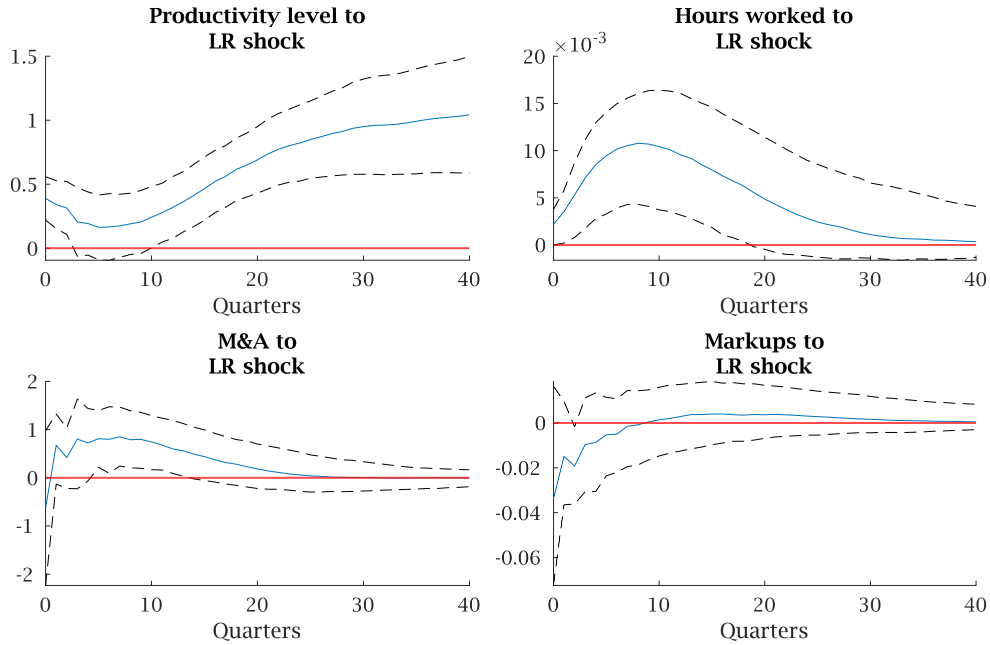


(b) Variance decomposition

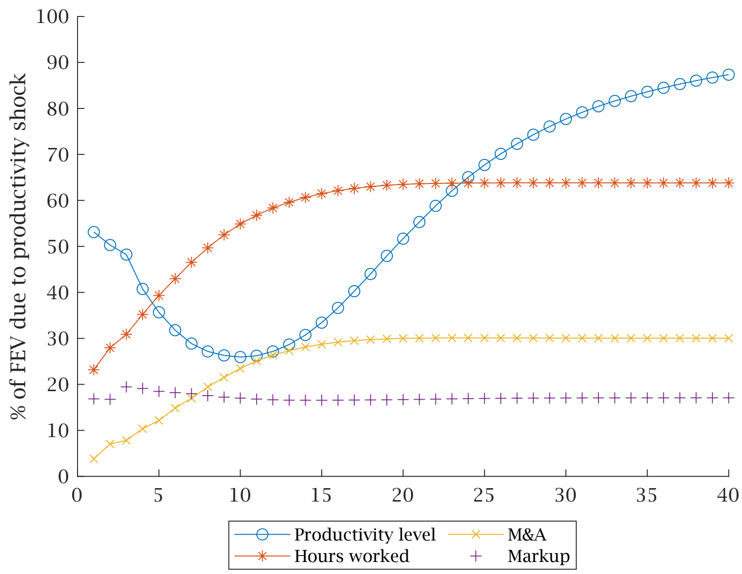
Figure 8: Results in case of $z_t = (apl_t, hw_t, m_t, \mu_t^{med})'$. Panel (a) plots impulse response functions to long-run productivity shock. Panel (b) plots corresponding variance decomposition.

tribution of markups respond to productivity shocks. Figure 9 (a) plots the impulse response functions. We see that responses of four variables are similar to the baseline case with weighted average markups. Finally, panel (b) of Figure 9 reports the associated variance decomposition. About 30% of forecast error in markups is associated with long-run productivity shock.

Figures 10 and 12 reports the full set of impulse responses under detrended median and 90th percentile time series estimates. Figures 11 and 13 report corresponding variance decompositions.



(a) Impulse responses on long-run productivity shock.



(b) Variance decomposition

Figure 9: Results in case of $z_t = (\Delta apl_t, hw_t, m_t, \mu_t^{90})'$. Panel (a) plots impulse response functions to long-run productivity shock. Panel (b) plots corresponding variance decomposition.

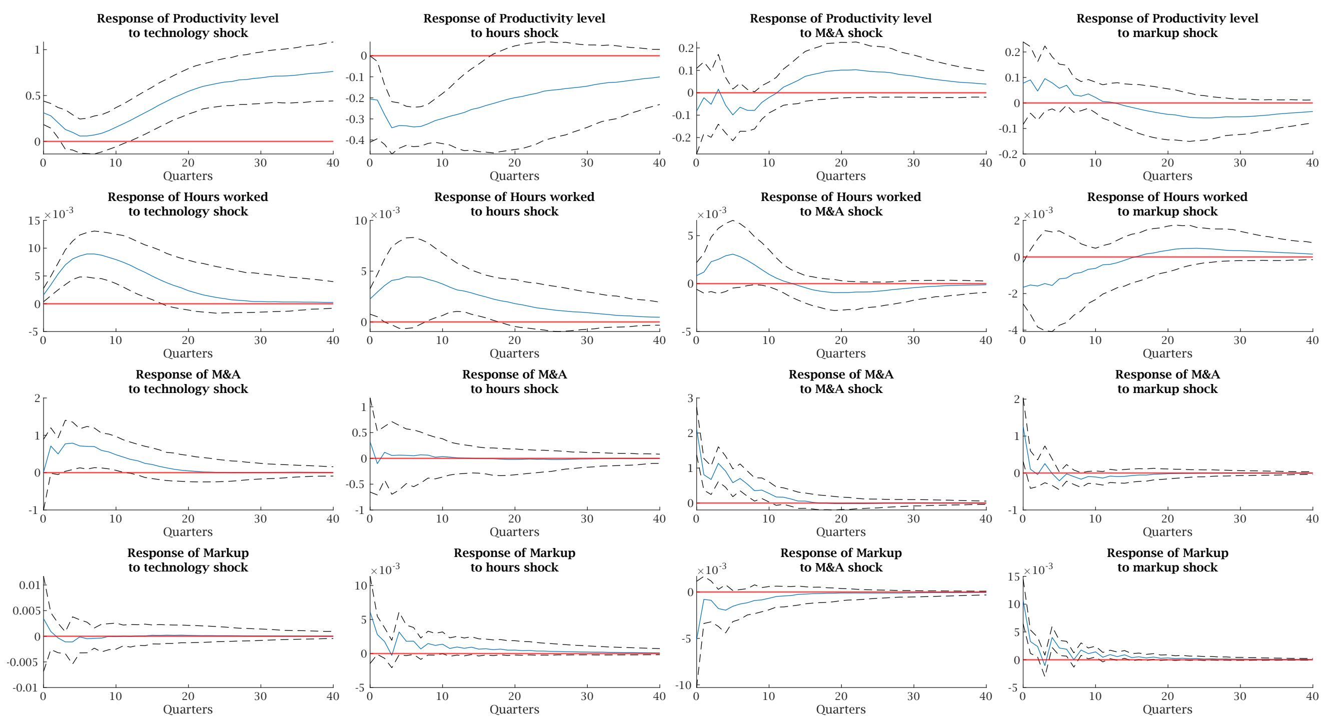


Figure 10: VAR(4) structural impulse response functions. $z_t = (\Delta apl_t, hw_t, m_t, \mu_t^{Med})'$

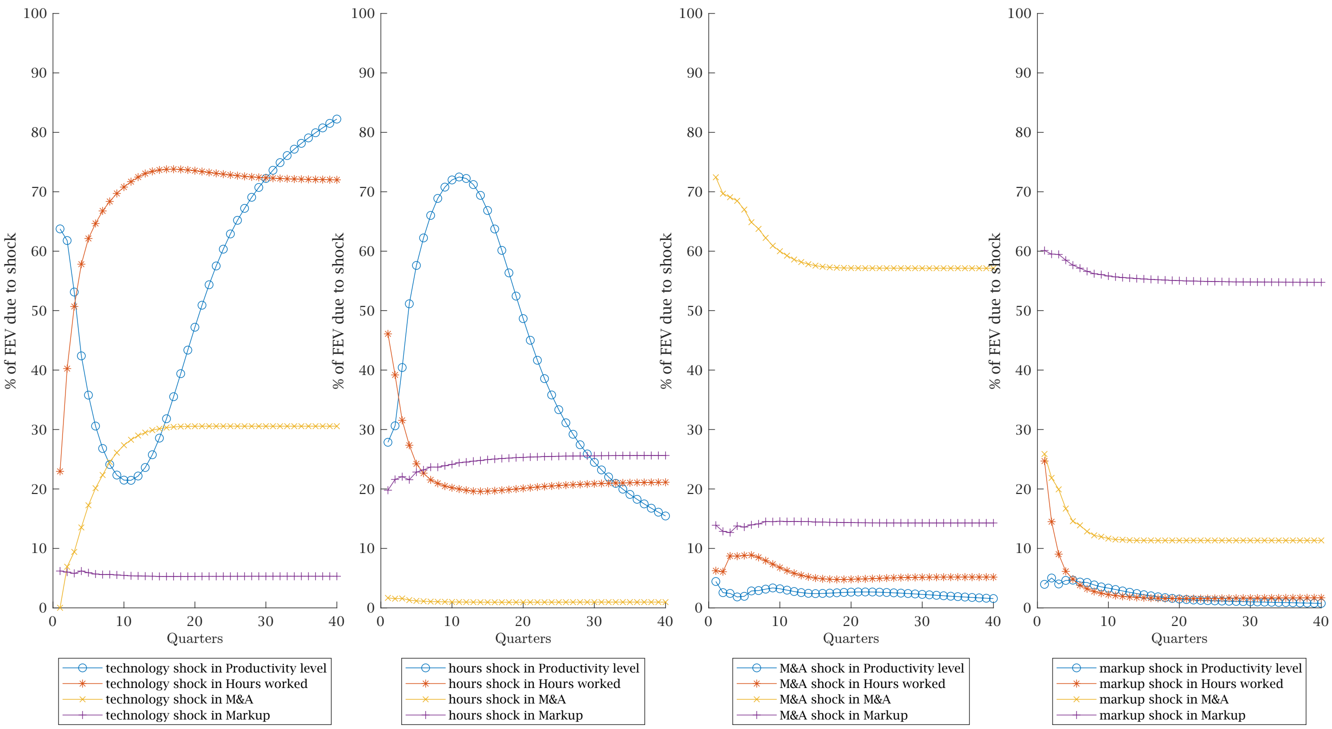


Figure 11: VAR(4) variance decompositions. $z_t = (\Delta apl_t, hw_t, m_t, \mu_t^{Median})'$

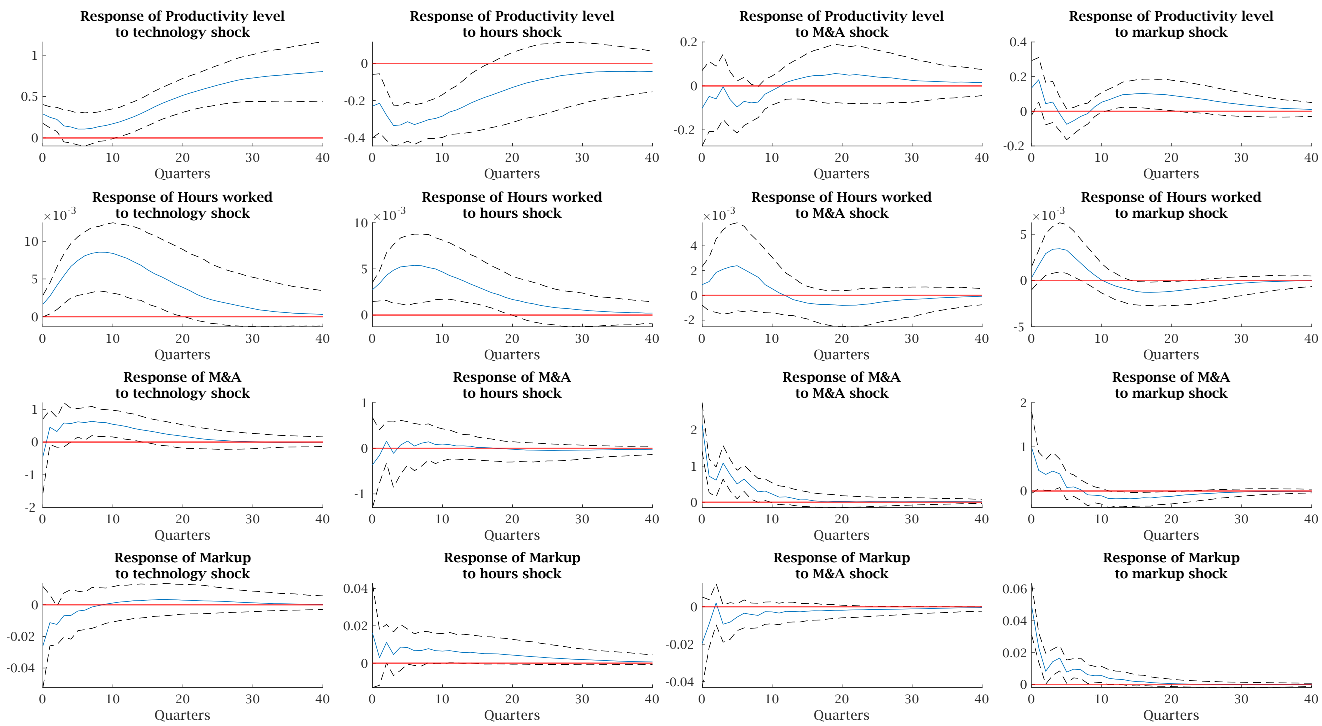


Figure 12: VAR(4) structural impulse response functions. $z_t = (\Delta apl_t, hw_t, m_t, \mu_t^{90p})'$

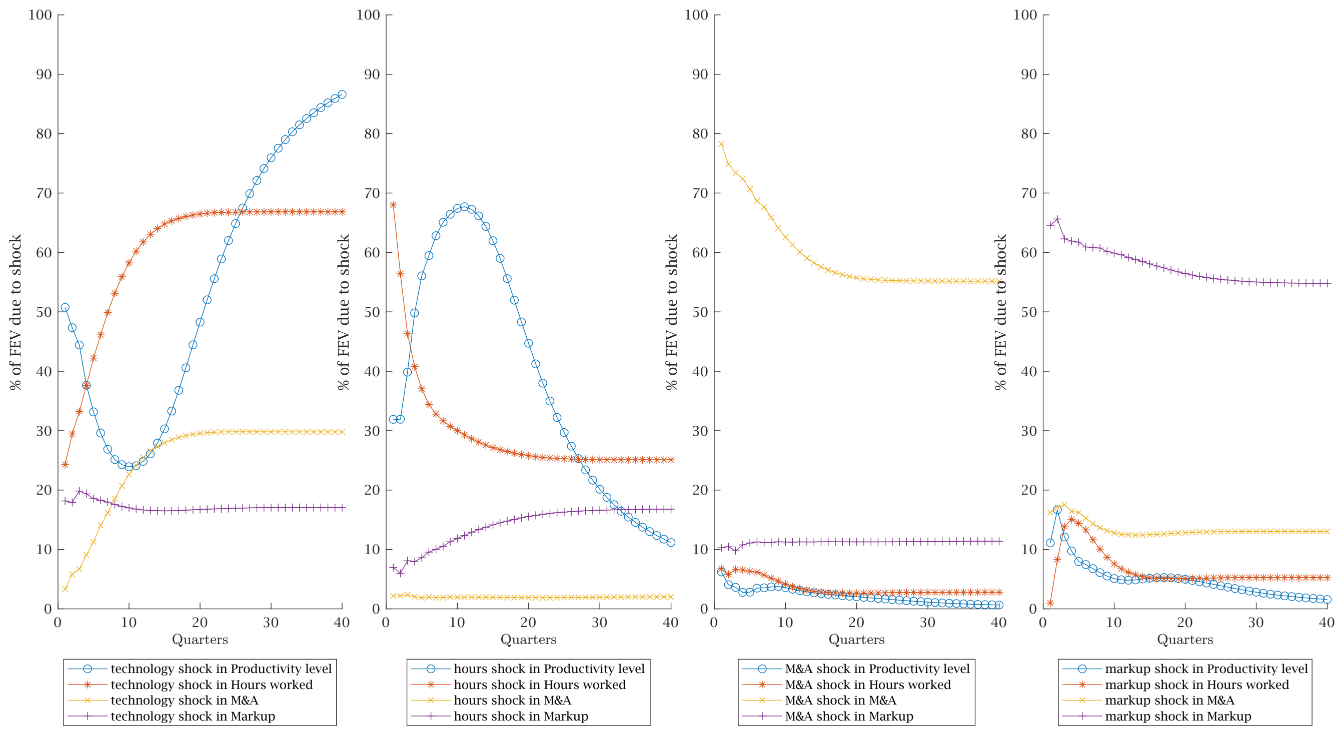


Figure 13: VAR(4) variance decompositions. $z_t = (\Delta apl_t, hw_t, m_t, \mu_t^{90p})'$

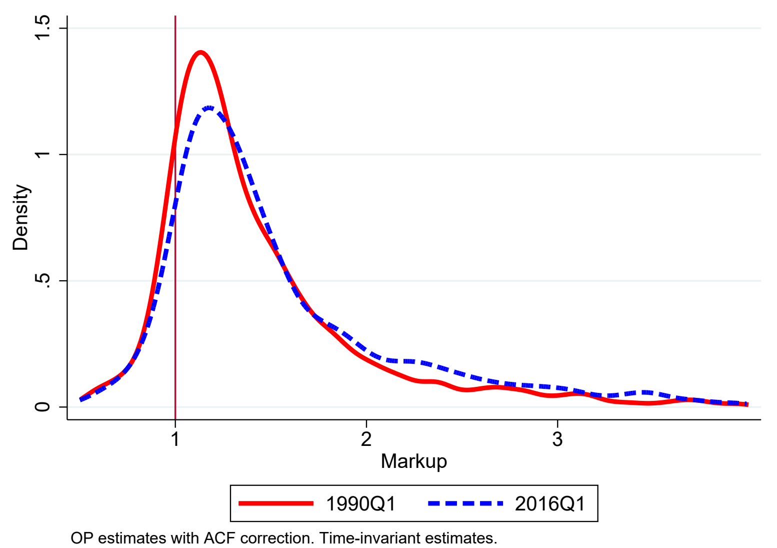


Figure 14: Distribution of firm-level markups. The firm-level markup estimates are obtained using Olley-Pakes type method with Ackerberg-Caves-Frazer correction.