UNITED NATIONS UNIVERSITY
UNU-WIDER

WIDER Working Paper 2019/58

Markups and market structure in South Africa

What can be learnt from new administrative data?

Joshua Budlender*

August 2019

United Nations University World Institute for Development Economics Research

wider.unu.edu

Abstract: The South African economy is generally understood to be characterised by high levels of product market concentration and high firm markups. This paper reviews the existing literature and discusses what can be learnt from new administrative firm-level panel data. I present new evidence on South African markups, industrial concentration, and the firm-size distribution, for sectors across the South African economy. I find that conclusions on whether markups are 'high' or 'low' are heavily dependent on the method used, and I show that this is consistent with the prior literature. There is however preliminary evidence that markups have generally declined over the 2010–14 period. I argue that it is difficult to make strong conclusions about industrial concentration using cross-industry study, and that high and growing concentration across the South African economy is yet to be conclusively shown. I also investigate how firm-level markups are related to industry-level concentration and firm-level market share. While some patterns emerge, I argue that their economic meaning is unclear.

Key words: markups, market power, concentration, firm-size distribution, South Africa

JEL classification: L11, D22, C81

Acknowledgements: I thank the National Treasury and UNU-WIDER teams whose assistance and support made this project possible. In this respect I owe a particular debt of gratitude to Amina Ebrahim, whose generous help and patient advice on subjects ranging from data problems to administrative issues went far beyond what could reasonably be expected. For comments and inputs at various stages of this project I thank seminar participants at UMass Amherst, participants of UNU-WIDER's Seminar Series and Inequalities Workstream meetings, Adam Aboobaker, Ihsaan Bassier, Aalia Cassim, Murray Leibbrandt, Jesse Naidoo, and Imraan Valodia. I thank James Traina for advice on markup es- timation methods, and Seidu Dauda (World Bank Group), Sara Nyman (World Bank Group) and Aalia Cassim (National Treasury) for Stata markup estimation code. All errors remain my own.

*University of Massachusetts, Amherst; jbudlender@umass.edu

This study has been prepared within the UNU-WIDER project Southern Africa - Towards Inclusive Economic Development (SA-TIED).

Information and requests: publications@wider.unu.edu

ISSN 1798-7237 ISBN 978-92-9256-692-0

https://doi.org/10.35188/UNU-WIDER/2019/692-0

Typescript prepared by the author.

The United Nations University World Institute for Development Economics Research provides economic analysis and policy advice with the aim of promoting sustainable and equitable development. The Institute began operations in 1985 in Helsinki, Finland, as the first research and training centre of the United Nations University. Today it is a unique blend of think tank, research institute, and UN agency—providing a range of services from policy advice to governments as well as freely available original research.

The Institute is funded through income from an endowment fund with additional contributions to its work programme from Finland, Sweden, and the United Kingdom as well as earmarked contributions for specific projects from a variety of donors.

Katajanokanlaituri 6 B, 00160 Helsinki, Finland

1 Introduction

The development of the new South African Revenue Services-National Treasury (SARS-NT) firm-level dataset of South African firms offers wide-ranging possibilities for novel work examining the firm-side determinants of important features of the South African economy. However because of the historic dearth of firm-level data of this type, large gaps remain in our knowledge of even basic features of this side of the economy. This paper seeks to address one such gap, by providing baseline descriptive evidence on market power and market structure across the South African economy. I review the existing literature and present original evidence on these questions. I conclude that consistent with the existing literature, the picture that emerges depends heavily on methodological choices which have mostly been insufficiently scrutinised. I include a detailed discussion about what can and cannot be concluded about these phenomena using the new data and prevailing methods, and suggest that some circumspection is required when using these approaches.

Substantively I present new evidence on the South African firm-size distribution, levels of industrial concentration, and on the distribution of firm-level markups. I then examine how these factors are related in a regression framework. I use both accounting and structurally estimated markups, and discuss the assumptions needed for these markups to be identified. In general I present careful discussions of the data and methodological issues associated with this type of study. This work falls broadly under the old Structure-Conduct-Performance (SCP) tradition in Industrial Organization (IO), and there are lessons to be learned from the international literature as to the strengths and weaknesses of this type of analysis. At the same time, in estimating markups I draw on a contemporary and highly active literature on production function estimation, and I discuss the costs and benefits of these structural methods.

I find that conclusions as to the levels of markups depend on methods used. While accounting markups in South Africa are relatively low, structurally estimated markups are high. I show that this disjuncture between accounting and estimated markups is consistent across the South African literature, and therefore that prevailing views that high markups characterise the South African economy are only supported to the extent that we believe the structural methods to be credible. I present a detailed discussion of the methods used to estimate markups in this paper, which I hope will be useful for the purposes of evaluating the estimates but also for other researchers who estimate production functions using the dataset. A rare consistent finding across methodologies is that while accounting and estimated markups are found to be very poorly correlated, they both show declining trends in markups between 2010 and 2014 when the dataset is restricted to a narrow sample of firms with relatively complete tax records.

My results suggest that the South African firm-size distribution is highly skewed, but that it seems to be comparable or slightly less skewed than is the case for other countries. Measured levels of industrial concentration are presented, but I discuss why it is difficult to make meaningful conclusions on this issue using cross-industry studies and the SARS-NT data in particular. I show that the existing South African cross-industry literature, which has suggested high and rising concentration in manufacturing industries, in fact presents evidence which is far from conclusive on these questions.

Divergences between methods and sub-samples are again evident in the regression analysis, but some consistent patterns can be found. In general it seems that estimated markups are frequently positively correlated with 4-digit industry market share, though when using accounting markups this correlation is significantly negative in the Wholesale and Retail sector. The correlation between markups and 4-digit industry concentration varies across sub-samples but is often negative. I argue that one should not put too much weight on these results, as they are unlikely to be informative about these phenomena in well-defined product markets. I find that the inclusion of a structurally estimated productivity proxy consistently reduces the partial correlation between markups and market share, but explain that this productivity proxy likely captures much more than output productivity and in fact may be correlated with price-setting power. In general the regression analysis is consistent with the overarching conclusion of this paper that cross-industry study may be too crude an approach to investigate diverse market structures and modes of competition.

I start by briefly reviewing the international IO literature on cross-industry study of market structure and market power, and highlight some important lessons to be learned. I then present an overview of the existing South African cross-industry literature, and argue that this literature does not provide robust evidence on the evolution of market structure and markups in South Africa. I proceed to the methods used in this paper, paying particular attention to the logic of structurally estimated markups. I show that these markups, and especially the productivity proxy which comes out of this approach, need to be considered with some circumspection. This note of caution is again emphasised in the data section, where I discuss the data and my adjustments in some detail. I then present summary statistics on the firm-size distribution, concentration, and markups, and end with regressions to show partial correlations. I conclude by relating my results to the existing South African literature, and argue that this collective evidence does not justify definitive conclusions about market structure and markups in the South African economy. I recommend that policymakers recognise the uncertainty we have about these phenomena, and suggest that further research interested in market structure, profits and modes of competition would be well-served by taking an approach which can incorporate greater market-level institutional detail.

2 Literature review

2.1 Markups and market structure in contemporary IO research

Investigations into the relationship between profitability and market structure have a long history in IO. Starting from Bain (1951, 1956), cross-industry empirical studies associated with the dominant mid-century Structure-Conduct-Performance (SCP) school were frequently used to investigate how markups were associated with industry concentration (Schmalensee 1989; Shepherd and Shepherd 2004). However as criticisms mounted regarding the theoretical foundations of SCP-style analysis, and key empirical results were overturned or proved difficult to verify, cross-industry study fell out of use in favour of careful industry-specific investigations (Davis and Garcés 2009). Today few IO economists study cross-industry trends in markups and concentration, but a resurgence of interest in these topics from economists in fields such as Macro and Labour Economics has led a recent spate of empirical works on these issues (Autor et al. 2017; Barkai 2017; Kehrig and Vincent 2018). Berry (2017) has argued that some of this new literature ignores well-founded lessons of the SCP demise – though he acknowledges the value of IO analysis which turns back to answering questions about "the economy", rather than "the market for yoghurt". A promising development in this area is the work of De Loecker et al. (2018), which uses modern IO methods to present evidence that economy-wide markups in the USA have been secularly increasing since the 1980s.[1]

A basic but fundamental critique of SCP-style analysis concerns the data. When doing cross-industry analysis, there is often little choice but to use accounting data on firm profits and costs. However since at least Fisher and McGowan (1983), IO economists have been skeptical about the degree to which accounting measures correspond to economic quantities of interest. The key issues relate to how accounting cost relate to economic costs, and what this means for the comparability of rates of return.

[1] A more well-known initial version of this article was circulated as De Loecker and Eeckhout (2017); De Loecker et al. (2018) is the latest version. De Loecker and Eeckhout (2018a) presents evidence on the change in markups globally using similar methods and data.

However there are also more basic issues when using reported accounting data, such as the problems caused by firms intentionally misreporting for tax or other strategic advantages. In the South African context, Wier and Reynolds (2018) present evidence that large South African firms significantly under-report profits by shifting incomes to tax havens.

Separate from these data issues is the cogency of the theoretical foundations of SCP analysis, and particularly investigation of the reduced form relationship between markups and concentration. For Bain (1951, 1956) and others, the generally positive relationship between markups and concentration was taken as empirical support for the exercise of market power through collusion. However as Berry (2017) recapitulates, in even the very basic undifferentiated product Cournot model, there will be a relationship between markups and concentration that does not indicate anything like collusion. In this model concentration is not an exogenous variable which can be perturbed to examine effects on markups, but is jointly determined with markups in equilibrium.

The critique about a lack of theoretical foundations to the SCP school ultimately became devastating. In a landmark paper Demsetz (1973) argued that the positive correlation between industry-level concentration and industry-level average mark-up could be completely unrelated to collusion or any kind of abuse of market power, and could instead be spuriously generated by firm-level market shares being positively correlated with firm-level mark-ups – where firms with high market share have low costs of production. That the relationship between concentration and markups was in fact a market share-markups relationship was confirmed by Ravenscraft (1983). Whether the market share-markups relationship was driven by efficiency (as per Demsetz (1973)) or by anti-competitive exercise of "relative dominance" (as per SCP defenders such as Shepherd (1972)) became an interminable debate, and its inconclusiveness contributed to the decline of SCP (Bresnahan 1989).

The descriptive value of cross-industry regression nonetheless remains. Schmalensee (1989) and Bresnahan (1989) in their respective handbook chapters discussing the decline of SCP and the rise of the "New Empirical Industrial Organization" both make clear that despite the general inability of cross-industry regression to establish causal relations, they can still uncover important empirical regularities which are essential guides to more causally-oriented research. However the modern use of cross-industry regressions must be cognisant of the lessons learnt since the SCP heyday, and Schmalensee (1989) emphasises the importance of recognising the descriptive rather than causal aims of this research, the use of robust specifications, and the use of multiple measures of key variables where possible. There is still much to learn from cross-industry study in a country such as South Africa, where basic evidence on market structure and markups does not exist – but this research needs to be done with these SCP critiques in mind. It is to the South African evidence on these questions to which I now turn.

2.2 The existing South African evidence

For the purposes of this paper the existing South African evidence on market structure and markups can be split into three themes: evidence on concentration, evidence on markups, and evidence on the relationship between these phenomena. All of these studies except Edwards and van de Winkel (2005) focused exclusively on South African manufacturing industries. Below I address each theme in turn.

Concentration

The first academic study of industrial concentration in South Africa was P. G. Du Plessis (1978), who used 1972 Census of Manufacturing data to construct various concentration measures at the 3-digit SIC level – and concluded that concentration in South African industries was very high. P. G. Du Plessis

(1978) was unusual, however, in that he had access to data which allowed calculation of standard absolute measures of concentration, such as concentration ratios (CRs). In much of the subsequent literature – such as Reekie (1984), Fourie and Smit (1989), Leach (1992), Fedderke and Szalontai (2009) and Fedderke and Naumann (2011) – researchers were forced to use relative concentration measures such as the Gini coefficient or 5% CR, or an inferior absolute measure such as an approximated Rosenbluth index.[2] Gini coefficients and Rosenbluth indices were calculated by approximating a Lorenz curve from the Census of Manufacturing data and then numerically integrating along the curve. In general researchers using these methods reported high concentration which was increasing over time up until 1996 (Fedderke and Szalontai 2009), but then reported a substantial drop in 2001 (Fedderke and Naumann 2011). Fedderke et al. (2018) present the latest research on concentration, using 2010-2012 firm level tax data, and report 3-digit industry relative concentration which is usually higher in the 2010-2012 period than was found for 1976-1996. They argue that the large break in the time-series of concentration in 2001 probably suggests that this data is not methodologically comparable to the rest of the series.

The standard story is therefore one of generally high and steadily increasing concentration, with perhaps a break in 2001. There are however some issues with the existing evidence. Firstly, as noted by Fedderke and Simbanegavi (2008) and Fedderke et al. (2018), the 2001 data likely cannot be compared to the rest of the series. While the pre-2001 and post-2001 data come from firm censuses or administrative records, the 2001 data comes from a survey. The dramatic drop in concentration between 1996 and 2001 – ostensibly a larger change in industry structure in five years than over the previous 20 years (Fedderke and Naumann 2011) – does not inspire confidence in the comparability of the data sources. The finding of Fedderke et al. (2018) that 2001 presents a serious aberration in the series reinforces this point. There is thus insufficient evidence to suggest an early 2000s decline in concentration.

It is also unclear, however, whether the conclusion of generally high and increasing concentration is well-supported. Fourie (1996) is unique in the literature in that he obtained access to 1980s Census of Manufacturing data which allowed computation of some standard absolute measures as well as relative measures of concentration. While he concludes that there is moderate evidence of increasing concentration over time, examination of his reported results for the absolute measures suggests relative stability rather than noticeable trends. Indeed as Fourie (1996) acknowledges, most differences in point estimates are not statistically significant. The exception is the Gini – the purely relative measure – where a noticeable increase is evident. And yet much of the South African evidence on increasing trends subsequent to Fourie (1996) is drawn from changes in the 5% CR – also purely relative. With the disjuncture between relative and (preferred) absolute measures evident in Fourie (1996), it is unclear whether absolute concentration has actually been high and increasing. Indeed when looking at over-time changes in the Rosenbluth index from Fedderke and Szalontai (2009, p. 244), 17 of the 24 manufacturing sectors they examine exhibit *declines* in concentration (unlike the Gini, which almost uniformly increases in the same time period). Fedderke and Simbanegavi (2008) make the additional point that the Rosenbluth numbers in the South African literature are all very low, indicating a competitive economy compared to US benchmarks. This may be due to the very broad market definition used by almost the entire South African literature – the 3-digit SIC industry – but as discussed in Section 4.2 this broad market structure just raises further questions about whether the reported concentration measures are at all meaningful in economic terms. Like Fedderke and Simbanegavi (2008), I conclude that the cross-industry evidence on high and rising concentration in South Africa is far from conclusive.

[2] Relative measures are generally seen as inferior to absolute measures for the purposes of measuring industrial concentration, because all else constant they are invariant in the number of firms in a given industry (Fedderke and Simbanegavi 2008). In this respect the 5% CR – the percentage of the market share controlled by the top 5% of firms – should be distinguished from standard *n*-firm CRs (which are absolute measures). The Rosenbluth index is generally seen as inferior to standard absolute measures such as the HHI because it is will react strongly to changes in the market shares of very small firms (Fedderke and Naumann 2011).

Markups

The "first wave" of papers on markups in South African industry – such as Reekie (1984), Leach (1992, 1997) and Fourie and Smith (2001) – were primarily concerned with the SCP-style markups-concentration relationship discussed below, and often did not report markup summary statistics. These papers calculated manufacturing accounting markups (discussed in Section 3.2). Where these were reported, markups were usually relatively low – in the range of 15% to 30% (Reekie 1984; Leach 1997). However the "second wave" of South African evidence on markups – mainly associated with Johannes Fedderke and coauthors – has in general *estimated* markups using Roeger's (1995) method of the "nominal Solow Residual" (NSR), based on Hall's (1988) relation between the residual and the mark-up. In general these estimated markups are found to be high – between 50% and 80% – with Fedderke et al. (2007) and Aghion et al. (2008) reporting markup levels approximately double those found for the United States. Aghion et al. (2008) and Fedderke and Hill (2011) find no secular trend in markups over the period they examine (1970-2004). Fedderke et al. (2018, p. 128) calculate accounting markups for 2010-2012, which though still quite high are generally lower than the previous estimates – and on this basis they conclude that "liberalising economic policies... may have put downward pressure on markups over time."

Is this evidence enough to conclude that South African manufacturing markups are generally high, and what can be said of trends over time? One issue which needs to be addressed is reconciliation of the first wave's low markups with the high markups of the second wave, even when the periods under consideration overlap. Methodology may provide an explanation: accounting methods may generally yield lower markups than structural estimation which attempts to get at marginal rather than average costs. Some evidence for this view is found in Aghion et al. (2008, p. 752), where estimated markups are high but accounting markups are low and in line with the numbers found by the first wave of South African research.[3] An implication of this explanation would be that the Fedderke et al. (2018) conclusion of declining markups may be premature. However even when looking only at estimated markups, the picture is not absolutely clear. Edwards and van de Winkel (2005) estimate lower manufacturing markups than Fedderke et al. (2007), and as Fedderke and Simbanegavi (2008) note, there is no clear explanation for the discrepancy. An issue may be lack of robustness of the estimation method, as is briefly mentioned by Aghion et al. (2008). New evidence from S. Du Plessis, Katzke, Gilbert, and Hart (2015) casts further doubt, with the authors using firm-level data from publicly traded firms to compare the profitability of the largest South African industrial firms to equivalent firms of comparator countries, and finding no evidence of higher South African profits. This divergence may be explained by S. Du Plessis et al. (2015) using profits rather than markups, or correcting profits for exchange rate depreciation (not an obviously desirable adjustment), but their findings remain striking.

Ultimately these issues do not overturn the general conclusion that markups seem to be high in South Africa, but it is worth noting that these high estimates come almost exclusively from structurally estimated models, and the finding is therefore only evident to the extent that the structural methods are plausible. This is an issue which has thus far received scant attention in the existing South African literature.

The concentration-markups relationship

As mentioned, the "first wave" of South African literature on markups was preoccupied with the SCP concentration-markups relationship. In particular, these authors were intent on distinguishing between

[3] Further evidence for this hypothesis is provided in this paper.

SCP and Demsetzian explanations for a raw positive correlation between markups and concentration which they all found at the industry level. The first study, Reekie (1984), used 1976 industry-level data from the Census of Manufacturing to reproduce Demsetz's (1973) analysis as best could be done with the sub-standard concentration data outlined above. Unlike Demsetz, Reekie (1984) found that concentration increased profits for both large and small firms, and accordingly concluded in favour of the SCP paradigm. Leach (1992) criticised Reekie primarily on the grounds of his handling of the concentration data, and produced his own analysis over a more extended time series supporting the Demsetzian view. This finding was recapitulated by Leach (1997) with the superior Fourie (1996) data, using methods as per Chappell and Cottle (1985). Fourie and Smith (2001) concluded this literature by noting methodological weaknesses of the cross-industry SCP methods, and producing careful findings which supported both Demsetzian and classical SCP explanations to varying degrees – a result they suggested was unsurprising.

In the "second wave" of the South African markup literature there are two exceptions to the general rule of avoiding the structure-profits question – Fedderke et al. (2007) and Fedderke et al. (2018). Fedderke et al. (2007) use a 1970-1997 industry-level panel based on Censuses of Manufacturing, and find that estimated markups are positively correlated with measures of industrial concentration developed by Fedderke and Szalontai (2009). However they only find an affect of concentration between industries, not within industries over time. Given endogeneity concerns regarding cross-sectional variation, the latter type of evidence would be more persuasive. Additionally the data used – a panel of 22 manufacturing industries over 28 years with 616 observations – seems to be heavily imputed. Their concentration figures come from Fedderke and Szalontai (2009), which is based on only 7 years of Census of Manufacturing data, and it seems they use linear interpolation to provide data for the 21 of their 28 time-periods not in Census of Manufacturing years (Fedderke and Szalontai 2009, p. 242; Fedderke and Naumann 2011, p. 2931).

Fedderke et al. (2018) use the first release of the SARS-NT firm-level panel. For each year of 2010-2012, they estimate a cross-sectional industry-level regression of the form $M_i = \beta_0 + \beta_1 C_i + \beta_2 A_i + \beta_3 A_i C_i + \varepsilon_i$, where M_i is the accounting markup, C_i is the concentration ratio, and A_i is average assets, all for 3-digit manufacturing industry i. They find that β_1 is a fairly precisely estimated 0, and that there is a large and statistically significant positive estimate for the interaction term β_3. These results suggest a positive correlation between markups and concentration for all industries with positive average assets (i.e. presumably all industries).[4] They then run industry-specific firm-level regressions, replacing concentration with market share, and find heterogeneous results for the analogous β_3 across industries. It is unclear what these coefficients indicate. While Fedderke et al. (2018) interpret the interaction as showing high concentration and barriers to entry, it could instead be understood as (efficiency-induced) market share from economies of scale.

Apart from specific issues outlined above, the Fedderke et al. (2007) and Fedderke et al. (2018) papers suffer from common weaknesses when it comes to determining the relationship between market power and markups. As the earlier review of the Demsetz-SCP debate explained, this kind of cross-sectional industry-level relationship does not establish abuse of market power. The conclusions in Fedderke et al. (2007, p. 54) – strong and unequivocal causal claims – are overstated given the evidence. As per Schmalensee (1989), robust descriptive evidence should be the objective. But even consistent descriptive evidence on the markup-concentration relationship is difficult to discern from the existing South African

[4] At one point Fedderke et al. (2018, p. 121) state that they "show markups are not correlated with industry concentration." This is not the correct interpretation of their concentration results, as calculating the marginal effect of concentration from their regression would show. At other points they conclude that there is no "direct" or "monotonic" relationship between concentration and profits (Fedderke et al. 2018, p. 139), but it is somewhat unclear what exactly that means in this context. I do not intend to quibble over semantics, but it is probably important to clarify what the results actually indicate – a positive markup-concentration correlation.

literature. One possible explanation is that it is data deficiencies which have obscured clear patterns. This paper explores whether clear patterns emerge from a new and more comprehensive dataset, which allows firm-level analysis across the South African economy. However another possibility is that the heterogeneity of markets and modes of competition across the economy is just too great to be usefully modeled by cross-industry study which cannot incorporate the intricacies and institutional details of individual markets – including such basic issues as establishing well-defined product markets. I will conclude in this paper that this latter explanation is very plausible.

3 Method

3.1 Overall approach

Given the issues outlined above, a major focus of this work is descriptive statistics. Under market structure I present summary statistics on the shape of the South African firm-size distribution, and evidence on the distribution and level of 4-digit industry concentration as measured by HHI. Markups are estimated as per Section 3.2, and summary statistics showing the levels and dispersion of markups are similarly provided. Multiple markup measures are presented and their distributions compared. For the main summary statistics, results are aggregated at the 1-, 2- and 3-digit industry levels.[5] In general, aggregated results are weighted by gross sales – though the weighting scheme will always be indicated. There is currently an open debate about how best to weight aggregate markups (De Loecker and Eeckhout 2018b; Edmond, Midrigan, and Xu 2018). As per De Loecker et al. (2018) I use gross sales as a natural weighting scheme, but it would be informative to compare these results with aggregate markups weighted by costs as recommended by Edmond et al. (2018).

The second part of the paper presents regression results, and investigates the relationship between markups and market structure. It is important to emphasise that this is also descriptive evidence – no claim to causal interpretation is made. For my main results I estimate a firm-level fixed effects specification of the form

$$\ln(\mu_{fit}) = \beta_1 m_{fit} + \beta_2 hhi_{it} + \beta_3 m_{fit} \cdot hhi_{it} + \gamma_1 Z_{fit} + I_f + I_t + u_{fit}, \tag{1}$$

where μ_{fit} is the firm-level markup of firm f in 4-digit industry i at time t, m_{fit} is the log of firm level market share at the 4-digit level, hhi_{it} is the log of the 4-digit HHI, Z_{fit} is a control which includes the log of the firm-level capital-sales ratio, I_f is a firm-level fixed effect and I_t is a year fixed effect. Equation 1 is estimated separately (by OLS) for each 1-digit industry, with cluster-robust standard errors clustered at the 4-digit industry level. As discussed in Section 4.2, I only present results for 1-digit industries where there are at least 20 4-digit sub-industries, as the cluster-robust variance estimator is unreliable if there are too few clusters (Cameron and Miller 2015). Despite using firm age as a control when estimating production functions as per section 3.2, I do not include it in Z_{fit} in equation 1. If high markups are sustained by barriers to entry, and firm age is correlated with barriers to entry (which seems plausible), then firm age would be a "bad control".[6]

For those markups which are estimated as per De Loecker and Warzynski (2012), it is possible to recover an estimate of firm productivity $\hat{\omega}_{fit}$ (see Section 3.2). I include $\hat{\omega}_{fit}$ as a covariate in some specifications

[5] See Section 4.2 for discussion of the industry variables used. 1-digit results are presented in the main paper, while 2-digit and 3-digit results are available in an online appendix. Some statistics are removed in order to preserve the anonymity of firms in the SARS-NT data; in every case it will be apparent where this has happened.

[6] Notwithstanding that these are descriptive rather than causally framed regressions, I do not want to include covariates which may control for the part of a structure-profit relationship which may indeed be caused by the exercise of market power.

to examine the extent to which the structure-profit relationship is attenuated by this control. As discussed in section 3.2, however, $\hat{\omega}_{fit}$ captures more than "productivity", so care must be taken when interpreting these regressions.

I do not specify a particularly rich model for markups in equation 1.[7] The objective of this paper is not to uncover all of the determinants of markups, but to investigate quite narrowly the SCP-style structure-profit relationship in the South African context, and to do so in a way which interrogates the robustness of the method, data, and results. In this respect I take to heart the recommendations of Schmalensee (1989) and Berry (2017) when it comes to the descriptive value of SCP-style analysis being in the establishment of a few key robust stylised facts.

3.2 Markup estimation

Markups are prices over marginals costs. The particular specification can vary, but common expressions in the literature are the price-cost ratio $\frac{P}{MC}$ and the Lerner Index $\frac{P-MC}{P}$. For comparability with the more recent international literature I use the price-cost ratio, but these quantities are clearly directly related. It is however not straightforward to observe markups in the data, because marginal cost is not an accounting quantity. The contemporary view is that marginal costs – and therefore markups – require some kind of structurally derived estimation.

Accounting markups

In this paper I use two broad approaches for measuring markups. The first is accounting markups, where the markup μ is given by:

$$\mu_{ft} = \frac{TR_{ft}}{TVC_{ft}} = \frac{TR_{ft}/Q_{ft}}{TVC/Q_{ft}} = \frac{AR_{ft}}{AVC_{ft}}. \tag{2}$$

TR_{ft}, TVC_{ft}, AR_{ft} and AVC_{ft} are respectively the total revenue, total variable costs, average revenue and average variable costs of firm firm f in time-period t, while Q_{ft} is its output. Even if we accept the costs reported in the data as being the economically relevant measures we are interested in, these accounting markups are generally not equivalent to the price-cost ratio defined above unless $AR = P$ and (more problematically) $AVC = MC$. Accounting markups nonetheless have a long history in IO, and are central to the early South African literature with which I seek to compare my results. As such I report accounting markups in addition to the estimated markups discussed below. Accounting markups tell us *something* of interest (even if they are not strictly markups over marginal cost), and this dataset presents a first chance to calculate these quantities across the South African economy.

Estimated markups

The contemporary IO approach to markup estimation is to use estimates from structural models. Usually, these approaches require specification of a model for firm conduct and detailed data on consumer demand, and as such have typically been restricted to market-specific studies (De Loecker et al. 2018). However for estimates of markups across a variety of markets, De Loecker and Warzynski (2012) offer an appealing and increasingly popular approach which does not require detailed industry-specific data and modeling. Instead, based on an assumption of cost-minimising firms, firm-level markups can be

[7] For a much richer analysis of the determinants of markups in the South African context see Dauda et al. (Forthcoming).

estimated using firm input and ouput data, plus a specified production function. I use the De Loecker and Warzynski (2012) approach for markup estimation in this paper. While I briefly summarise some features of the approach here, I present a more detailed explanation in Appendix A.

Estimation.—The key insight (derived in Appendix A) is that the markup $\mu_{ft} = \frac{P}{MC}$ can be implicitly expressed via

$$\frac{\partial Q_{ft}(\cdot)}{\partial X^v_{ft}} \frac{X^v_{ft}}{Q_{ft}} = \mu_{ft} \frac{P^{X^v}_{ft} X^v_{ft}}{P_{ft} Q_{ft}}, \tag{3}$$

where a firm f produces output Q_{ft} in time t according to the production technology $Q_{ft}(\cdot)$. It uses V variable inputs X^v (such as labour and materials) in addition to capital. $P^{X^v}_{ft}$ is the input price of variable input X^v while P_{ft} is the price of output Q_{ft}. Simplifying notation, I denote the output elasticity with respect to a variable input X as $\theta^X_{ft} \equiv \frac{\partial Q_{ft}(\cdot)}{\partial X_{ft}} \frac{X_{ft}}{Q_{ft}}$ and the expenditure share of variable input X in total revenue as $\alpha_{ft} \equiv \frac{P^X_{ft} X_{ft}}{P_{ft} Q_{ft}}$. The variable input is typically taken to be labour, and indeed this is what I use for all of my reported specifications. The markup can be expressed as

$$\mu_{ft} = \theta^X_{ft} \frac{P_{ft} Q_{ft}}{P^X_{ft} X_{ft}} = \theta^X_{ft} (\alpha^X_{ft})^{-1}. \tag{4}$$

The intuition behind this result is that in the absence of markups (i.e. $\mu = 1$), a cost-minimizing firm will equate its expenditure share on variable input X to the output elasticity of that variable input. This can be quite clearly seen in equation 3. The markup drives a wedge between the expenditure share of a variable input and its output elasticity, and as per equation 4 it can be identified by the extent of this wedge.[8] There are of course a number of assumptions underlying this approach, discussed below. However these issues notwithstanding, the central challenge when implementing this approach is estimating the quantity θ^X_{ft}.

The input's revenue share α_{ft} can (subject to measurement error) be readily observed in the data. However the output elasticity θ^X_{ft} requires the specification and estimation of a production function, a fundamentally difficult exercise. There is a substantial literature on production function estimation, with the seminal contributions being Olley and Pakes (1996), Levinsohn and Petrin (2003) and Ackerberg et al. (2015). De Loecker and Warzynski (2012) implement an approach essentially the same as Ackerberg et al. (2015) (henceforth ACF), which I follow.

In my main value-added translog specification I seek to estimate a production function of the form

$$y_{ft} = \beta_l l_{ft} + \beta_k k_{ft} + \beta_{ll} l^2_{ft} + \beta_{kk} k^2_{ft} + \beta_{lk} l_{ft} k_{ft} + \omega_{ft} + \varepsilon_{ft}. \tag{5}$$

Lower-case letters indicate logs, so that y is observed logged output, l is the log of labour, and k is the log of capital stock. ε_{ft} is an unanticipated shock to production, while ω_{ft} is productivity unobserved by the econometrician (but potentially known to the firm). Given that I use labour as my "free" variable input, I am interested in estimating the output elasticity $\frac{\partial y_{ft}}{\partial l_{ft}} = \beta_l + 2\beta_{ll} l_{ft} + \beta_{lk} k_{ft}$. The problems of simply estimating equation 5 by OLS are well-recognised, however, and are primarily due to the unobserved productivity term. At the heart of the ACF approach (drawing on Olley and Pakes (1996)) is the introduction of a proxy variable for productivity, which allows identification of the coefficients in equation 5 via a two-step control function procedure. The key assumption is that material inputs are (conditionally) strictly increasing in productivity, which allows an inversion of the materials function to derive some

[8] Another way to think about this relationship is to multiply both sides of equation 3 by X^v_{ft}/Q_{ft} – and we are then left with an expression where if markups equal 1, the marginal product of the variable input equals its real wage (ie the marginal cost to the firm). Markups are thus identified by the wedge between the marginal product and marginal cost of a variable input to production.

expression for productivity $h_f(\cdot)$ which is in turn a function of inputs to production, including materials, and some control variables (see Appendix A equations 10 and 11).

The innovation of ACF is to recognise that with $h_f(\cdot)$ by necessity being estimated non-parametrically (as per equations 12 and 13, Appendix A) in the "first stage", the coefficients of interest in equation 5 require additional structure in order for them to be identified. To this end they impose a law of motion for productivity, with an AR(1) process being the common choice. This allows estimation of the "second stage" where productivity is projected on its lag and coefficients can be identified using standard GMM techniques and some additional assumptions to provide moment conditions (see equations 14 and 15 in Appendix A).

With an estimate for the output elasticity of labour now in hand, it is straightforward to calculate the markup as per equation 4.[9] However while the approach outlined above yields a firm-specific markup, the production function needs to be estimated over some group of firms. Following De Loecker and Eeckhout (2018a), I implement the production function estimation separately per two-digit industry, for each industry pooling all 5 years of my data.[10] The translog specification means that output elasticities will still vary by firm and time period, due to heterogeneous l_{ft} and k_{ft}, but the underlying production technology is invariant within 2-digit industries and across time. A last note worth mentioning is that the De Loecker and Warzynski (2012) method is also able to yield an estimate for ω_{ft}, productivity, which comes out as a Solow-type residual (see Appendix A). As discussed below some care must be taken when interpreting this term.

Remarks.—There is nothing novel in the methodology outlined above and in Appendix A. I follow very closely the methodologies outlined in De Loecker and Warzynski (2012) and ACF, and indeed for the implementation of the production function estimation above I use an already-existing Stata program, *prodest*, from Rovigatti and Mollisi (2016).[11] Production function estimation is however an active area of research, and there are some increasingly well-known problems with the De Loecker and Warzynski (2012) and ACF procedures. An intended contribution of this paper is to outline some of these for an audience which may be interested in using the SARS-NT data to investigate market power.

A first issue concerns adjustment costs in labour. When using labour as the freely adjustable input, the De Loecker and Warzynski (2012) markup is identified by the wedge between the output elasticity of labour and its revenue share, with the assumption being that this wedge can only be caused by markup pricing as per equation 3. However if there are notable dynamic adjustment costs for labour inputs – such as fixed term contracts or labour law protections – this wedge may also be affected by an inability of the firm to optimally adjust in the short-run. In this case estimated markups may be substantially biased (up or down) in particularly rigid labour markets.[12] If one uses a gross output production function and materials are assumed to be freely adjustable, markups can in principal be estimated by looking at the wedge between the output elasticity and revenue share of materials.[13] There are some problems with this solution, however. Most seriously, Gandhi, Navarro, and Rivers (2018) argue that gross output

[9] One additional detail is that instead of calculating the expenditure share α^X_{ft} directly from the data, I use the estimate of ε_{ft} obtained from the first stage regression to purge the observed output of measurement error – see Appendix A.

[10] See Section 4.2 for details on industry classification.

[11] *prodest* is frequently updated. I use the most up-to-date version as of January 2019.

[12] This may seem to be a particularly severe problem in the South African case, where the labour market is often taken to be highly rigid. This assumption may not be well-founded, however, with Kerr (2018) showing that South African worker flows and churning are substantial.

[13] This requires amendment of the moment conditions in equation 15 to include moments in materials, and if one makes the strong assumption that the labour market is so rigid that current labour is uncorrelated with the current-period productivity shock, then current-period labour can be used as an additional moment (Ackerberg et al. 2015).

functions are not identified under the ACF procedure. Practically, and as discussed in Sections 4.2 and 4.3, there are major problems with the identification of material inputs in the SARS-NT data and markups estimated using materials exhibit extreme volatility.

Issues with the SARS-NT measure of material inputs lead to a second concern, which is the crucial assumption that material inputs are (conditionally) strictly increasing in productivity. It is this monotonicity condition which allows the inversion of the materials function and the formation of a productivity proxy (see equation 10 and equation 11 in Appendix A). De Loecker and Warzynski (2012) note that a large class of models predict that materials will be increasing in productivity, but I know of no work examining this fundamentally empirical issue in South Africa. A still more troubling feature of the approach to productivity is the specification of an AR(1) process for productivity – a restrictive functional form assumption which implies that the productivity process is completely exogenous. While in principal this can be relaxed (De Loecker and Warzynski 2012), in much of the applied work a simple AR(1) process is used (De Loecker et al. 2018; Traina 2018).

While other issues remain (such as the assumption of time-invariant industry-wide production technologies and perfectly competitive input markets), I lastly focus on potentially major problems caused by observing output as *revenue* rather than physical product. While all of the theory outlined above was on the basis of output and inputs being in physical product terms, it is much more often the case that at least output and capital are observed in revenue or expenditure terms – and this is the case with the SARS-NT data. De Loecker and Warzynski (2012) suggest that price heterogeneity will cause output elasticities (and therefore markups) to likely be downward-biased, while De Loecker and Goldberg (2014) note that there can be bias in either direction. Whatever the direction of the bias, it is a problem. Apart from issues in the estimated markup levels, markups between sectors will be biased to different degrees, depending on industry-specific price heterogeneity. This is an unsolved problem in the production function and markup estimation literature – and it seems that practitioners have generally adopted the practice of noting the issue but proceeding with their analysis nonetheless. A notable exception is De Loecker et al. (2018), where they attempt to address the issue by controlling for markups in their production function estimation, using a proxy approach where markups are taken to be functions of industry-year linear time trends and firm-level market share. In this paper I include market share as a control in the proxy function $h_t(\cdot)$, and it is hoped that this will absorb some of the variation in the revenue term y_{ft} due to markups. The issue should however not be skirted: current approaches will yield biased estimates for markups, and results are therefore necessarily approximations.

The issue of observing revenues rather than physical output has particularly severe consequences for the use of $\hat{\omega}_{ft}$ as a productivity proxy.[14] As equations 5 and 14 (in Appendix A) make clear, productivity $\hat{\omega}_{ft}$ is estimated as a residual. When output is in physical terms, this residual may quite plausibly be understood as productivity. However when output is revenue, it will also represent market power or other pricing phenomena which are correlated with the arguments of $h_t(\cdot)$. It is quite plausible that input demand will be correlated with output prices, and it is likely that pricing will be correlated with controls I use in the productivity proxy function, such as market share and firm age. Therefore while part of $h_t(\cdot)$ may indeed be controlling for productivity, in my view it is incorrect to view $\hat{\omega}_{ft}$ as a simple measure of productivity, and in some applications this may lead to severe problems where productivity is conflated with pricing power.

Structural approaches come with both benefits and costs. The purpose of the preceding section was to lay out the markup estimation method transparently and with critical discussion, so that researchers interested in South Africa can decide how much weight to put on the markups presented here. My own view is that the theoretical concerns are severe enough – especially when considered in tandem with data

[14]It is natural to want to use the ACF method to examine determinants and correlates of productivity. Kreuser and Newman (2018) is recent paper which does exactly this using the SARS-NT data.

problems outlined below – that the evidence presented using these methods must be seen as suggestive rather than simply taken at face-value.

4 Data

The data used for this paper is a new South African firm-level panel constituted from tax records. The dataset is the result of a partnership between the South African Revenue Services (SARS), the National Treasury (NT), and the United Nations University World Institute for Development Economics Research (UNU-WIDER). It is selectively made available to researchers at the NT offices in Pretoria, South Africa, and access is granted on the basis of strict non-disclosure rules and the requirement that results do not allow identification of individual firms or persons. The panel is a merged dataset from four different sources of South African tax data: Company Income Tax (CIT) records, Customs records, Value-added tax (VAT) records, and worker-level Pay-as-you-earn (PAYE) records from IRP5 forms. The latter three sets of records are matched to single CIT-registered entities, but not all records in each dataset are fully merged. The CIT records are the backbone of the dataset I use here and contain balance sheet and income statement information. I also make significant use of the IRP5 data aggregated to the CIT-firm level. For a much more detailed discussion of the dataset see Pieterse et al. (2018).

The SARS-NT panel is still very new, and is frequently updated.[15] Because the data is still a work-in-progress, and the emphasis has been on updating it and developing facilities for data access, it is also not particularly well documented. Its novelty, and the general novelty of firm-level data in South Africa means that there is still no established best practice when it comes to using the data, and researchers frequently use quite different approaches to data cleaning or spend significant time exploring the data to uncover issues which other researchers have already discovered and corrected for in their own analysis. I therefore spend some time discussing the dataset and data cleaning choices I make, both so that this methodology is reproducible and also to contribute towards the development of some standard practices when it comes to the data. Below I frequently refer to variable names as they are in the data.

4.1 Dataset restrictions

In the main merged CIT panel there are 6 426 631 firm-year observations, spread over 8 years from 2008 to 2015.[16] Given concerns about the reliability of data in 2008 and 2009 (Pieterse et al. 2018), and that there are only about 200 000 firms in 2015 compared to between 800 000 and 900 000 firms per year otherwise, I restrict my analysis to firms present between 2010 and 2014. The panel is unbalanced, as firms enter and exit the dataset, and becomes more unbalanced as I drop some firm-year observations as part of the cleaning process. Because it is data necessary for most of the analysis presented here, and also as part of an effort to identify "real" firms (rather than dormant firms, shell companies, or other non-productive tax vehicles), I drop firms which have non-positive or missing information on gross sales (`g_sales`), fixed capital (`k_fixed`), or total variable costs, where these are defined as the sum of cost of sales (`g_cos`) and labour expenses (`x_labcost`)[17] from the firm income statement. I also drop

[15] I use the versions of the CIT and IRP5 datasets as they were in July 2018. At the time of writing (March 2019) the dataset had just been updated from the July 2018 version, but with limited time at the data centre I could not re-run my analysis from the new base data.

[16] Because I work with IRP5 data matched to the CIT, I use the `taxyear` rather than `finyear` variable for years – see Pieterse et al. (2018).

[17] I adjust the labour expenses variable where it seems that it inconsistently does not include directors income.

firms indicated as dormant, and those which list their industry as "Public administration".[18] This results in a substantial reduction in firm counts, from 4 414 045 firm-year observations over 2010-2014 to 1 080 234 firm-year observations over the same period. When I restrict this to firms which have non-missing data for my industry classification of choice (discussed below), I am left with 973 053 firm-year observations.

From this main dataset of 1 080 234 observations, which I call the "Broad" sample, I also create a more restricted "Narrow" set of firms. The Narrow sample is the Broad sample but excluding firms which i) are not matched with the IRP5 data, ii) have non-positive or missing employment (`irp5_empl_weight`), iii) have non-positive or missing materials costs, or iv) have non-positive or missing value-added.[19] This results in a significant number of firms being dropped – leaving me with 455 415 firm-year observations – which is overwhelmingly driven by the requirements that the CIT is matched by the IRP5 data and has positive non-missing employment information. When I restrict the Narrow sample to firms on which I have data for my preferred industry classification, I am left with 420 166 firm-year observations. On the one hand the use of the Narrow sample is necessitated by some of the techniques used below, and in order to meaningfully compare techniques I need to estimate across comparable datasets – so the Narrow sample is an unavoidable reality. On the other hand, it is perhaps undesirable that there are substantial numbers of firms in the Broad sample which have non-positive employees or materials costs, and it may be desirable to use the Narrow sample as it further identifies "real" firms, or at least firms which have reasonably accurate and complete tax records. I report all main results for both the Broad and Narrow samples.

4.2 Variables

Broadly speaking two factors need to be considered when cleaning variables in the SARS-NT panel. The first is that it is clear that variables are reported with substantial error. It is very common for the largest few values of a continuous variable to be orders of magnitude larger than surrounding observations, and for this reason all continuous variables used in the analysis are trimmed at the 1st and 99th percentiles of their yearly distributions. It is more difficult when dealing with error in categorical variables, as the discussion of industries below makes clear. The second issue which needs to be considered arises because the dataset was not created for researchers: it is administrative tax data. As such, many variables are not exactly what a researcher requires, and alternative measures need to be constructed. I discuss this below in the section on constructing measures of firm wage bills and materials costs. All variables expressed in nominal prices are adjusted to be in real 2014 prices, using the overall South African CPI.[20]

Industry classification

Industry classification is particularly important for a study such as this one, where results are frequently reported and estimated by industry. Firms report their industry in 3 different places in the tax records upon which the SARS-NT panel is built.[21] In the CIT forms (IT14 and ITR14), they report in two

[18] See below for discussion of the industry classification system I use.

[19] Materials costs, value-added and the labour bill are all derived variables. I explain their creation below.

[20] While certain producer price indices are reported in the data, they are for 2-digit ISIC PAYE industry classifications. I do not use this industry specification and thus approximate price changes with the overall CPI adjustment.

[21] I ignore VAT industry codes also available in the data, as this variable is extremely poorly populated.

places: firms use a SARS "profit code" – a type of industry classification – to indicate their "main source of income", and on the same page report the "source code" – which is meant to be an SIC code – for their "main industry". In the IRP5 data, each worker has a "main income source code", which is an industry classification similar to the CIT SARS profit code. The SARS codes are transformed into ISIC codes by an NT-SARS crosswalk, and in the dataset firms can be classified into industries in three ways: the 5-digit SIC code from the CIT forms (the SIC industry code), a 3-digit ISIC code from from the CIT profit code (the ISIC profit code), and a 3-digit ISIC code from IRP5 main income source code (the ISIC PAYE code).[22] While ISIC and SIC are slightly different classification schemes, all codes are cross-walked into the same 17-sector classification, which I call the 1-digit level.

While the harmonisation of the SIC and ISIC codes into the 17-sector 1-digit level should be relatively straightforward, the different 1-digit industry variables in fact match very poorly. To give some examples: of the firms in the Broad sample classified in the 1-digit categories of "Mining and quarrying", "Electricity, Gas & Water" and "Information & Communications" for the SIC industry codes, only 52%, 28%, and 19% respectively are classified into the same 1-digit industry per the ISIC profit codes. The matching is worse for the Narrow dataset. In particular it seems that far more firms are classified into manufacturing for the ISIC profit codes than is the case for the SIC industry codes. While there may be some legitimate differences in the data (if a firm for some reason views its "main industry" as different from the industry in which it makes most of its income), this level of disagreement cannot be accurate. It is unclear whether the error comes from the SARS-NT crosswalks (the origins of which are not documented), confusion from businesses as to how to fill in the form, or something else.[23] The issue may be exacerbated by industry imputation, but the poor matching exists in the un-imputed data too, as far as I can see.[24]

I ultimately choose to use the SIC industry code as my preferred classification for 3 reasons. The main reason is that it allows greater industry disaggregation than the ISIC codes. While the ISIC codes only disaggregate to the 3-digit level, the SIC industry code disaggregates to the 5-digit level. I use the 4-digit disaggregation, as most reported 5-digit codes are in fact 4-digit level classifications. As discussed immediately below, I need as finely disaggregated industries as possible. Secondly, the SIC industry code is less reliant on the SARS-NT cross-walks than the ISIC codes, and given the lack of documentation I prefer to avoid the crosswalks where I can. Lastly, the SIC industry codes are preferable to the ISIC PAYE codes as they cover far more firms in the Broad dataset.[25] There is however clearly more work to be done in terms of understanding what drives the industry classification discrepancies, and which is the best variable to use.

Before continuing it is worth being explicit about the problems with using 4-digit industries for concentration and market share statistics, regardless of which industry classification scheme is used. While it was standard to use 4-digit industries to define markets in the older SCP literature, and De Loecker et al. (2018) do the same in their contemporary analysis, these are in general *not* well-defined product markets. It is not the case that all firms within a specific 4-digit industry will be competing with each other or producing for the same set of buyers. Firstly there is a problem of geographic scope. For these 4-digit industries to be well-defined product markets, it would need to be the case that all firms produce

[22] It seems that the modal worker-level IRP5 code is used to create a CIT-level IRP5 ISIC industry code.

[23] Brief discussion with some SARS officials suggests that the issue may be in the underlying reporting. SARS will be attempting to restrict the possibility of inconsistent reporting for the 2019 tax submissions.

[24] The imputation of industry codes in the SARS-NT panel is not documented, but a coding error which prevented the imputation from implementing allowed a look at the data prior to that particular procedure (there may be further imputing prior even to this stage, but this unlikely). I re-implement the imputation, which iteratively assigns industry codes to firms which have missing industry data in one year if they have industry data in an adjacent year.

[25] The ISIC profit code covers a very similar but slightly higher number of firms than the SIC industry codes; and the ISIC PAYE codes cover slightly more firms than the SIC industry code in the Narrow dataset.

and compete nationally. While this will be the case for some firms, in others the relevant market for concentration and market share would be local, and the 4-digit industry will be incorrect. Secondly, the 4-digit industry is often too coarse a classification for product markets.[26] 4-digit industries such as "growing of fibre crops", "manufacture of plastic products" and "manufacture of consumer electronics" clearly will include firms operating in different product markets serving different sets of buyers. The economic relevance of concentration and market share in these large sectors is not obvious. Lastly, and related to the immediately preceding point, the number and size of 4-digit industry markets will depend significantly on the seemingly arbitrary degree to which a 1-digit sector is disaggregated in the SIC codes at the 4-digit level. For example while the "construction" and "mining and quarrying" sectors are divided into 8 and 14 4-digit sub-industries respectively, "manufacturing" is disaggregated into 136 such sub-industries. Concentration will therefore seem spuriously low in "construction" and "mining and quarrying" relative to "manufacturing" simply because the former sub-industries are defined more broadly. It is these issues which lead me to conclude that the SARS-NT dataset as it currently exists is in fact *not* well-suited for answering questions about economy-wide product market concentration, at least using the straightforward approach I take here.

For the regressions analysis I focus attention on 7 1-digit sectors which have more than 20 4-digit sub-industries, in order to have a credible minimum number of clusters. These industries are "Agriculture, Forestry & Fishery", "Manufacturing", Wholesale & Retail", "Transport & Storage", "Information & Communication", "Finance, Insurance & Real Estate" (FIRE), and "Administration & Support activities". Though it contains more than 20 sub-industries, I exclude "Other services" as a focus of analysis due to it's indistinct definition. The selected sectors are generally self-explanatory, with perhaps the exception of Administrative & Support services. Amongst other sub-industries this sector includes rental and leasing services, activities of employment placement agencies, private security, facilities support (including cleaning), labour broking in general, and other "human resources" provision.

Creating wagebill and materials variables

Apart from information on sales, capital, and labour, the De Loecker and Warzynski (2012) markup method also requires data on firm wage bills and on material inputs. The main source of information on firm expenditure and incomes in the SARS-NT data is the income statement in the CIT data, and it is natural to expect to use the labour expenses and cost of sales variables as measures of the wage bill and materials respectively. This is however not strictly correct. The cost of sales measure reported in the income statement, analogous to "cost of good sold (COGS)" reported in other contexts, may include labour costs and strictly speaking often *should* include labour costs. *Direct* labour costs – that is the cost of labour which is directly involved in the production of the good/service sold – should generally be included in cost of sales, while the wages of workers more removed from direct production or sales can be recorded under labour expenses. With this being said, firms have considerable leeway in how these quantities are reported, and it may be that reported cost of sales and labour expenses will quite often approximate materials and labour costs.

For my preferred specification I build a wage bill variable from the IRP5 data, and use this to back out a materials cost measure from the CIT income statement data. For each CIT entity per year, I sum up employee incomes recorded in the IRP5, excluding both directors' incomes and non-wage labour

[26]While the table is too large to reproduce here, in the online appendix I include a spreadsheet which details the industry classification scheme I use at the 1-digit, 2-digit, 3-digit, and 4-digit levels. It also includes the number of firm-year observations per 4-digit industry for the broad and narrow samples.

expenses such as share options, but including items such as overtime pay and special payments.[27] After trimming, I then compare this derived wage bill with the trimmed labour expenses variable, after taking directors' incomes (`x_director` from the CIT income statement) out of the labour expenses measure too.[28] In cases where the IRP5 wage bill variable is larger than the labour expenses variable (about 40% of the firm-year observations), I assume that this is because some labour costs are included in the firm's cost of sales, and subtract this difference out of (trimmed) cost of sales. I take the remaining cost of sales balance as a measure of material input costs (I drop the few firms with negative materials costs at this point). Perhaps somewhat remarkably, the distribution of the resultant cost of material inputs is ultimately very similar to the distribution of trimmed cost of sales reported in the CIT income statement. This is because the wage bill is generally quite similar to labour expenses, at least when director's income is removed. The aggregate similarity may of course hide substantial changes for individual firms. In any case, and notwithstanding their aggregate similarity, I prefer the derived wage bill because it relates the IRP5-derived labour variable I use to IRP5 data on the cost of that labour – and it seems preferable to use the same dataset for these quantities when markups are identified by the wedge between the output elasticity of labour and its revenue share.

4.3 Markup specifications

I present estimates for three different markups. In all cases I trim inputs at the 1st and 99th percentiles, and also trim the resultant markups at these levels.

The first set of markups are calculated as per the accounting approach in equation 2, for both the broad and narrow datasets. I use gross sales as the measure of total revenue, and the sum of cost of sales and labour costs (`x_labcost` from the income statement) as the measure of total variable costs.

My second, main specification of estimated markups (which I refer to as DLW markups with reference to De Loecker and Warzynski (2012)) are estimated according to equation 4. I calculate a Value-Added output term by subtracting my derived material costs term from gross sales, and use the same derived material costs term as the materials proxy for estimation. I use labour inputs as the variable input over which markups are estimated, and specifically use the number of full-time equivalent workers from the IRP5 data (`irp5_empl_weight`).[29] I use the SARS-NT measure of fixed capital costs (`k_fixed`) as the measure of capital, and my derived wage bill to calculate the input revenue share. Given the data requirements these markups can only be estimated for the narrow sample.

For my third specification, which is mainly a robustness check, markups are also estimated as per equation 4, but I use the income statement information on costs rather than the IRP5 data. Specifically I use labour expenses (`x_labcost`) as the measure of labour and of the wage bill, and cost of sales as the measure of materials. I call these markups Income statement markups. The purpose of this specification

[27]Prior to summing up these incomes, I follow the steps suggested in Pieterse et al. (2018) which are intended to remove repeat submissions and records not related to individuals.

[28]Micro-enterprises which submitted ITR14 (as opposed to IT14) CIT forms do not report directors' incomes separately from their total labour costs, and therefore I leave these labour expenses unadjusted. It seems that this does not cause much of a distortion because a) the IRP5-derived wage bills including and excluding directors' income do not differ by much for micro-enterprises and because b) few micro-enterprises are matched with the IRP5 data in any case, lessening their importance for this analysis.

I remove directors' incomes for two reasons. Firstly, directors' incomes as reported in the IRP5 and CIT match extremely poorly for some reason – they are generally substantially larger in the CIT. Secondly, my measure of employment is based on the number of employees reporting wage income (as discussed below). The wage bill should reflect the cost of this variable input, and as such it makes little sense to include the cost of directors in the wage bill.

[29]See Pieterse et al. (2018) for details on the construction of this measure.

is to check the robustness of the main DLW markup estimates for the narrow sample, but also to see how estimated markups change from the broad to narrow samples. Because the Income statement approach does not require IRP5 data, it can be implemented for the broad sample – though the requirements of positive cost of sales and labour expenses does mean the sample is actually narrower than the broad sample used to estimate accounting markups (which only require a positive *sum* of cost of sales and labour expenses).

All of the estimation approaches above use value-added translog production functions. I believe this is prudent given the concerns about gross output functions raised by ACF and Gandhi et al. (2018). I do however estimate a variety of other specifications. Though I do not report these results, I discuss them briefly in Appendix B.

5 Descriptive results

5.1 Market structure

Firm-size distribution

The international literature shows that the firm-size distribution is typically highly skewed. In a seminal paper, Cabral and Mata (2003) argue that while publicly available data suggests firm sizes are lognormally distributed, for more complete data the right-skew is even greater than lognormal – and they demonstrate this with the case of Portugal. They also show that firm size distributions are typically more highly-skewed in young firm cohorts, and converge to lognormal over time. Here I do not discuss firm growth, but present simple cross-sectional statistics.[30] I calculate firm-size statistics for both the broad and narrow samples, and find that the South African firm-size distribution is in fact quite closely approximated by a lognormal distribution. This is a somewhat surprising result given that South Africa is sometimes supposed to have an unusually skew firm-size distribution, and as Tybout (2000) discusses for the case of manufacturing firms, developing countries may in general be expected to have particularly highly-skewed (or bimodal) firm-size distributions.

For firm-size I use gross sales and 4-digit industry market share, and also use total employment for the narrow sample where I have this data. Figures 1a and 1b show these distributions estimated using kernel density estimation, and as can be seen they are typically very close to the overlaid log-normal density. The distribution of sales appear slightly more right-skewed than log-normal, but the market-share distribution is very similar. The distribution of firm employment seems to show bunching at low levels. Examination of industry-specific distributions (presented in Appendix C) suggest this is due to over-representation of firms with 1 employee.[31]

Dramatic differences between the broad and narrow samples are not evident from the figures, but differences can be seen from the industry-specific summary statistics presented in Tables 1a and 1b.[32] Apart

[30] For evidence on firm growth using the South African data, see Mamburu (2017).

[31] Industry-specific and year-specific kernel densities for the distribution of sales and market share are available on request. They do not show particularly notable inter-industry or inter-year heterogeneity.

[32] In these tables I use Gini coefficients as measures of dispersion, firstly because Axtell (2006) argues that second moments may not be well-defined for highly-skewed firm-size distributions, and secondly because it is a natural inequality measure, which is of some interest.

from notable industry heterogeneity, these tables show that sales and market share are typically higher but more equally distributed in the narrow sample. This reflects in the figures as longer left-hand tails in the broad distribution. The Gini coefficients are all quite high, especially in the broad distribution, showing substantial inequality in firm sales, market share, and employment. Per industry sales in the broad sample are more unequally distributed than what is usually found for South African household income, while in the narrow sample the Gini coefficients are comparable. Mean market share per industry is generally very low – typically less than 1%.[33] Unsurprisingly given the skewness of the data, there are large differences between mean and median employment for all industries.

Concentration

Tables 2a and 2b show statistics on 4-digit industry concentration, for each 1-digit sector. The dramatic differences between weighted and unweighted statistics are notable – indicating that the sub-industries with highest concentration are usually relatively small in terms of their total sales. While weighted statistics are probably preferable when trying to get an idea of the typical level of concentration in a 1-digit sector, the issue is not obvious. It is likely that overly broadly-defined 4-digit industries discussed in Section 4.2, such as "manufacture of plastic products", will constitute large shares of the total sales in their 1-digit sector, and therefore will significantly influence weighted means and medians. These issues also make it difficult to conclude meaningfully on the relative concentration of sectors. While Manufacturing is consistently reported as one of the most concentrated sectors in the broad sample, its 137 sub-industries mean the market definition used in Manufacturing is quite different to what is used for other sectors. For example while there has been significant regulatory action in South Africa regarding price-fixing by a construction cartel, the Construction sector comes out in Tables 2a and 2b as one of the least concentrated sectors - likely because the sector is disaggregated into only 12 4-digit sub-industries.

In order to more transparently examine concentration in the context of these issues, Figures 2a and 2b show changes in concentration over time for each 4-digit industry, by sector (restricting attention to the 7 sectors used in the regression analysis). These graphs make clear the degree of heterogeneity within some sectors compared to others, as well as showing whether high average concentration comes from a few relatively concentrated sectors or one outlier. For example while the FIRE and Transport & Storage sectors look quite similar in the aggregate statistics in Table 2a, the Transport & Storage sector has a number of highly concentrated sub-sectors, whereas within FIRE there is just one highly concentrated industry. These kinds of patterns do change somewhat when looking at the narrow sample. In general concentration increases, which is not surprising given that the narrow sample is comprised of fewer firms. There are however also some distributional changes, such as greater dispersion and volatility over time. Concentration in the FIRE sector is particularly changed, as a disproportionately large share of the firms excluded from the narrow sample are in FIRE (see Tables 3a and 3b).

Figures 3a and 3b summarise the information in Figures 2a and 2b by collapsing the sub-industries into sector level means and medians. The main takeaway from these figures is that there does not seem to be any notable time trend in concentration at the 4-digit industry level. The figures also show that the Agricultural, Forestry & Fishery, Manufacturing, and Administration & Support Services sectors are quite consistently the most concentrated sectors, but this conclusion is plagued by the sub-industry classification issues already discussed. Perhaps a more meaningful finding is that the Retail & Wholesale sector is made up of low-concentration sub-industries, despite having the second-largest number of 4-digit sub-industries. This of course does not preclude high concentration in more meaningfully defined

[33]This should not be taken as evidence of lack of market dominance. Apart from the fact that these 1-digit results are aggregated over many sub-industries, as discussed in Section 4.2 these sub-industries are inordinately large.

product markets, such as Von Broembsen (2016) describes for the South African supermarket sector – which again highlights the issue of market definition.

5.2 Markups

Per sector, pooled over time

Figures 4a and 4b show the (unweighted) pooled distribution of markups across all sectors and years. The markup distribution is clearly skewed for all markup distributions, as comparison to the overlaid lognormal densities makes evident. It is notable however that the accounting markups are particularly highly right-skewed, while the DLW labour markups are very close to being lognormally distributed. The "Income statement" markups over labour costs exhibit very long *left* tails in logs.

Tables 3a, 3b, and 4 pool markups across years per 1-digit sector. While mean accounting markups are extremely high in the broad sample, medians and the markups in the narrow dataset are typically lower and much more comparable to each other – and these comparable markups in the narrow dataset are almost all lower than DLW labour markups using the same data. Looking at the narrow dataset, the differences are quite stark. While the accounting markups suggest markups which are slightly lower than what De Loecker et al. (2018) find for the US, the estimated DLW labour markups (which are more comparable to De Loecker et al.'s (2018) method) are substantially larger than US markups. Mean DLW markups are particularly extreme, probably implausibly so in the case of the Wholesale and Retail sector at least. It is worth noting that in earlier versions of this paper the DLW Wholesale and Retail sector markup was large, but not in the implausible range suggested by these results. This likely speaks to the volatility of the markup estimation method implemented here.[34] Interestingly, the accounting markups presented here are quite similar to what De Loecker and Eeckhout (2018a) estimate for South Africa using Compustat data – and which they classify as medium-low markups. The DLW labour markups – mean or median – would put South Africa well into the highest category of global markups as per De Loecker and Eeckhout (2018a).

Weighted and unweighted markups are fairly similar across the accounting and DLW labour methods, with unweighted markups consistently slightly larger than weighted markups using the accounting method. This suggests that in aggregate, larger firms have slightly lower accounting markups.[35] For DLW labour markups weighting sometimes increases and sometimes decreases markup levels, depending on the industry.

The accounting and DLW markups also differ when it comes volatility or dispersion. The accounting markups in the broad sample have extremely high standard deviations, but this is no doubt related to the skewness of the data which similarly yields very high means. When looking at the narrow sample, accounting markup standard deviations are much more reasonable, and comparison across this same sample shows that it is the estimated DLW labour markups which exhibit substantial heterogeneity.

[34]While it is in principal possible that the divergence between the current and previous Wholesale and Retail markups are due to small data-cleaning changes undertaken as the project developed, this seems unlikely. Rather, as Mollisi and Rovigatti (2017) show, the ACF procedure is highly sensitive to the different starting points used in its numerical optimization routine, and results can be dramatically biased under any given implementation of the procedure.

[35]This aggregate finding does not mean that large firms *within well-defined sub-industries* have lower markups than their smaller rivals. It may be the case that large firms have *relatively* higher markups than other firms competing in their same industry, but that industries which have smaller firms have generally higher markups across the board. This is indeed what De Loecker and Eeckhout (2017) find for the US.

Markups estimated using Income statement data are shown in Tables 5a and 5b, and the levels of these markups vary substantially between industries. Also like the DLW labour markups, very high standard deviations suggest that there is substantial markup heterogeneity. This heterogeneity is however understated, because (as discussed in Appendix B) negative markups constitute a non-negligible portion of raw markups estimated using this approach, and negative markups are dropped from the sample. Additionally, the Income statement markups have changed dramatically across different versions of this project. In earlier iterations they were consistently higher than DLW markups, while in these final results the comparison varies substantially per industry. This again highlights the issue of estimation sensitivity discussed in footnote 34.

The Income statement markups are presented here mainly for robustness purposes and two conclusions become evident from their examination. Firstly, accounting markups are generally lower in the narrow sample than broad sample, and this also holds for Income statement markups as we move between samples. Note, however, that this did not hold in previous versions of this paper. Secondly, across the 3 measures there is substantial disagreement even in the relative ordering of sectors by markup. There is substantial markup heterogeneity and it is hard to draw any but the most basic general conclusions.

If focus is placed on the manufacturing sector markups, a striking conclusion which emerges is that the results are quite consistent with the existing South African literature (despite a new markup estimation method and different dataset). The accounting markups in the narrow dataset and the medians in the broad dataset generally fall into the 15%-30% range identified earlier as the bounds of the accounting markups in the early South African literature, while the estimated DLW markups are close to the 50%-80% range in the "second wave" of estimated South African markups, though means reported here are somewhat higher. This seems to confirm the importance of markup methodology, and that conclusions that South African markups are high will depend on the credibility of the estimation methods. While the estimated markups in the earlier literature do not use the same approach, the Roeger (1995) and De Loecker and Warzynski (2012) methods are both based on Hall's (1988) relation, and they thus both require similar (and strong) assumptions, as discussed. I return to the issue of markup credibility in the conclusion.

Comparatively and over time

Figure 5 presents a more direct look at how the different markup measures compare. The top panel shows the distribution of estimated markups along the accounting markup distribution, while the bottom panel shows accounting and income statement markups along the DLW labour markup distribution. Each point indicates one of 50 quantiles along the x-axis, also therefore giving some indication of the shape of the x-axis markup distribution.

What is immediately apparent from the figures is that the DLW labour markups and accounting markups are not strongly correlated. Surprisingly, there is some suggestion of a negative correlation (and indeed this is what is found when the raw correlation is calculated), with DLW labour markups being highest when accounting markups are low. In contrast, the Income statement markups tend to increase with the accounting markups. It is also clear (by looking at the plotted points in the top panel) that DLW labour markups are typically larger than accounting markups, except for the few highest quantiles of accounting markups. The bottom panel shows that the Income statement markups are also not highly correlated with the DLW labour markups.

Figures 6a, 6b and 6c plot 1-digit mean and median markups over time. As was evident from Tables 3a and 3b, mean accounting markups in the broad sample are inordinately large. However across all accounting markup specifications, markups are *relatively* high in the FIRE and Agriculture, Forestry

& Fishery sectors, and low in Wholesale and Retail. In the narrow sample, for both accounting and DLW labour markups, Information and Communication markups are relatively high. The most dramatic differences between the accounting and DLW labour markups concern the Wholesale & Retail markups. While these are low for accounting markups, the opposite is true for the DLW labour markups. As already discussed, the extremely large values of the Wholesale & Retail DLW markups are likely not plausible, but they have been consistently large across previous versions of this paper.

In the broad sample there is no obvious time trend in accounting markups, but there is a consistent decline in the narrow sample. This downwards trend in the narrow sample is even more dramatic for the DLW markups, which is quite striking given the poor correlation between accounting and DLW labour markups. This common decline is an interesting and novel finding, but further corroboration is needed before it is taken as strong evidence of declining markups in South Africa. On the one hand an immediate issue is that this is only evident in the narrow sample. Is this because it is a peculiar sample, with some kind of non-random attrition bias? Or is the narrow sample a better reflection of "real firms", with credible data? Another issue regarding the DLW labour markups is that the production function is estimated over all years of the data. For examination of time trends it would be better to estimate production functions separately by year, but then sample size issues may become a problem. De Loecker and Eeckhout (2018a) discuss how markups have generally increased for developed countries in recent years, while in developing countries they have often remained constant or declined. This may be part of that general phenomenon, or could be related to stagnant growth in South Africa over the period of the dataset. More evidence is required on these questions before any firm conclusions can be reached.

6 The structure-profit relationship

I now turn to fixed effect regressions as outlined in Section 3.1.[36] As mentioned previously, I only present results for the 7 industries which have at least 20 4-digit sub-industries (in order to have enough clusters), and exclude the ill-defined "Other services" sector. The firm-level fixed effects regressions used here control for time-invariant characteristics of firms, and the only identifying variation comes from changes in characteristics over time for a particular firm. Given that I use a non-linear specification (I use interacted variables), I focus on marginal effects for interpretation, though coefficients are reported.

6.1 The canonical model

A few general conclusions emerge from the regressions presented in Tables 6 and 7. For the accounting markups, across the broad and narrow datasets I find that market share is negatively correlated and concentration positively correlated with markups for the Wholesale & Retail sector, while the opposite is true for the Agriculture, Forestry and Fishery sector. For the narrow sample, across the accounting and DLW labour markups, market share is positively correlated with markups while markups decrease in concentration for the Agriculture, Forestry and Fishery; Manufacturing; and ICT sectors. This pattern of a positive market share correlation and negative concentration correlation holds across most industries for the DLW markups. When they are significant, effect sizes are often somewhat larger for the DLW markups than is found for accounting markups.

[36]While I follow the structure-profit nomenclature used in the literature for these types of regressions, this is somewhat of a misnomer. Markups are in general *not* equivalent to profits, both because of general issues regarding inference of economic profits from accounting data, and also because markups are explicitly about prices over only *variable costs*.

The results which find significant negative impacts of concentration and market share on markups are intriguing, and somewhat difficult to explain. One possibility for the negative market share effect – which could make some sense for the accounting markup results – is that it may indicate a sector where large market share firms pursue a high-volume low-margin selling strategy. Negative concentration effects could perhaps be related to firm exit – where concentration increases because adverse conditions cause some firms to exit, and the same adverse conditions decrease markups in the sector. Positive market share effects are more easy to hypothesize about. The classic Demsetzian explanation is that it indicates high efficiency firms achieving large market share and markups, while the Shepherd (1972) explanation would be that large firms exert relative market power over their smaller rivals. The old SCP test for distinguishing between these two explanations would be to look at the market share-by-concentration interaction. In cases where it is significantly positive, this is often taken as evidence of the relative dominance argument.

However I would argue for some circumspection before making these types of conclusions. Firstly, because the industry definition does not align with well-defined product markets, market share is likely more related to something like "firm size" rather than being a quantity which is truly related to dominance of a product market. For the same reason, it is very unclear what the economic meaning is of the broadly defined HHI quantities. As emphasized by Berry (2017), these types of regressions are fraught with confounding relationships even if one ignores product market definition issues, and the lack of clarity on what is actually being measured only increases my hesitancy in too over-confident an interpretation of these results. Additionally, as well-worn criticisms of hypothesis testing have made clear, when presenting multiple point estimates (as I do here), spuriously significant results will be found, and consistent effects across different specifications present more convincing evidence than any single regression coefficient. The issues with any particular coefficient's standard error are exacerbated by the problems of few clusters, which will tend to bias these standard errors downward (Cameron and Miller 2015).

6.2 With the productivity proxy

The introduction of the derived $\hat{\omega}_{fit}$ productivity proxy into the DLW regressions significantly increases the explanatory power of the regressions, with the adjusted-within R-squared increasing notably for most specifications. As would be expected given that the productivity term is by definition a residual explaining firm revenue product, it is positively related to markups for all sectors. It also consistently decreases market share coefficients, sometimes rendering them statistically significant and negative, while concentration effects generally become more positive. The change in market share effect is to be expected. If the productivity proxy captures real output productivity, the change in market share partial correlation is consistent with a Demsetzian efficiency interpretation, while if it captures market power then the productivity term may be a "bad control" which partials out the anticompetitive effects of relative dominance. The concentration change is more intriguing, and suggests that concentration is consistently negatively correlated with firm revenue product productivity.

7 Conclusion

In this paper I have presented some novel descriptive results concerning markups and market structure in South Africa. At the same time I have tried to be transparent about the limitations of the evidence presented here, which I have connected to weaknesses in the South African cross-industry literature in general. I find that markups are generally high when they are structurally estimated, and low when accounting markups are used – a result which is consistent with most of the South African literature.

This suggests that the conclusion that South Africa is characterised by high markups will depend on the credibility of the structural methods. I discuss the assumptions underlying the markups estimated in this paper, and show that they are quite restrictive. Additionally, I explain that the markup estimation method is quite volatile. This implies that high markups should be seen as suggestive rather than definitive. This should not be interpreted as saying that markups are probably low in South Africa. While I have not focused much discussion on the accounting markups, this is because their deficiencies are more transparent: they are markups over observed average variable costs, not marginal costs. The existing evidence does not definitively establish high or low markups across the South African economy.

If the high estimated markups over marginal cost and low accounting markups over average cost are taken at face value, one way to reconcile these results is to infer that average costs may be notably higher than marginal costs for South African firms. This could be explained by large fixed costs, or by significant unexploited economies of scale. Based on the volatility of the results presented here, the restrictive assumptions, and the data issues discussed, I do not advise that too serious a weight be placed on this explanation. But the relevant implication is that even if one wishes to accept the markups presented here, they still present more nuanced implications for competition policy than the simple story of high and rising markups from the existing literature.

An even greater agnosticism must apply to industrial concentration when using cross-industry approaches. I have argued repeatedly that concentration at the 4-digit industry is unlikely to be informative about concentration in well-defined product markets, which is the real quantity of interest for most purposes. This, incidentally, may also present further issues for the markups presented here, which are estimated at the 2-digit industry level. If unusually heterogeneous firms make up these two-digit industries, the assumption of a common production function would be particularly restrictive. Even ignoring this issue, I have argued that the existing South African evidence does not present strong evidence for high and rising concentration at the 3-digit level, despite this being commonly presumed. Again, I should not be interpreted as suggesting that concentration is low or is not rising – the conclusion I draw is that cross-industry study does not seem to get us very far in resolving these questions.

Ultimately, few stylised facts emerge from this exercise. Those that do emerge are new evidence – for example the skewness of the firm size distribution, the positive markup-market share relationship, or declining markups between 2010-2014 – but the greater contribution of this paper is to demonstrate the contingency of the empirical results which come out of the application of these methods, and some explanation of why the results may be fragile. My view is that policymakers should recognise that cross-industry study provides little robust evidence about markups and market structure in South Africa. A clearer picture will likely only develop as further work is undertaken which can take into account more industry-specific institutional detail. It is hoped that this paper may provide an entryway for this kind of work, by giving some approximate industry-specific results. At the very least this paper might prove useful in highlighting some of the pitfalls of cross-industry work for those who seek to draw on the existing evidence or do work in this tradition in the future.

References

Ackerberg, D. A., Caves, K., and Frazer, G. (2015). 'Identification properties of recent production function estimators'. *Econometrica*, *83*(6): 2411–2451.

Aghion, P., Braun, M., and Fedderke, J. (2008). 'Competition and productivity growth in South Africa'. *Economics of Transition*, *16*(4): 741–768.

Autor, D., Dorn, D., Katz, L. F., Patterson, C., and Van Reenen, J. (2017). 'The fall of the labor share and the rise of superstar firms'. NBER Working Paper No. 23396. National Bureau of Economic Research.

Axtell, R. (2006). 'Firm sizes: facts, formulae, fables and fantasies'. CSED Working Paper No. 44. The Brookings Institution.

Bain, J. S. (1951). 'Relation of profit rate to industry concentration: American manufacturing, 1936–1940'. *The Quarterly Journal of Economics*, *65*(3): 293–324.

Bain, J. S. (1956). *Barriers to new competition* (Vol. 3). Cambridge, MA: Harvard University Press.

Barkai, S. (2017). *Declining labor and capital shares.* (Unpublished paper (JMP))

Berry, S. (2017, November). 'Market structure and competition, redux'. In *Proceedings of the 10th annual ftc microeconomics conference.* Washington, DC: Federal Trade Commission.

Bresnahan, T. (1989). 'Empirical studies of industries with market power'. In R. Schmalensee and R. Willig (eds), *Handbook of industrial organization* (Vol. 2, p. 1011-1057). Amsterdam: North Holland.

Cabral, L., and Mata, J. (2003). 'On the evolution of the firm size distribution: Facts and theory'. *American Economic Review*, *93*(4): 1075–1090.

Cameron, A. C., and Miller, D. L. (2015). 'A practitioner's guide to cluster-robust inference'. *Journal of Human Resources*, *50*(2): 317–372.

Chappell, W. F., and Cottle, R. L. (1985). 'Sources of concentration-related profits'. *Southern Economic Journal*,(): 1031–1037.

Dauda, S., Nyman, S., and Cassim, A. (Forthcoming). 'Product market competition, productivity, and jobs: The case of South Africa'. World Bank Policy Research Working Paper. World Bank Group.

Davis, P., and Garcés, E. (2009). *Quantitative techniques for competition and antitrust analysis*. Princeton: Princeton University Press.

De Loecker, J., and Eeckhout, J. (2017). 'The rise of market power and the macroeconomic implications'. NBER Working Paper No. 23687. National Bureau of Economic Research.

De Loecker, J., and Eeckhout, J. (2018a). 'Global market power'. NBER Working Paper No. 24768. National Bureau of Economic Research.

De Loecker, J., and Eeckhout, J. (2018b). *Some thougts on the debate about (aggregate) markup measurement.* (Unpublished paper)

De Loecker, J., Eeckhout, J., and Unger, G. (2018). *The rise of market power and the macroeconomic implications.* (Unpublished paper)

De Loecker, J., and Goldberg, P. K. (2014). 'Firm performance in a global market'. *Annual Review of Economics*, *6*(1): 201–227.

De Loecker, J., and Warzynski, F. (2012). 'Markups and firm-level export status'. *American Economic Review*, *102*(6): 2437–71.

Demsetz, H. (1973). 'Industry structure, market rivalry, and public policy'. *The Journal of Law and Economics*, *16*(1): 1–9.

Du Plessis, P. G. (1978). 'Concentration of economic power in the South African manufacturing industry'. *South African Journal of Economics*, *46*(3): 172–182.

Du Plessis, S., Katzke, N., Gilbert, E., and Hart, C. (2015). 'Mark-ups and competition: a comparison of the profitability of listed South African industrial companies'. Stellenbosch Economic Working Paper No. 02/15. Economics Department and Bureau of Economic Research, Stellenbosch University.

Edmond, C., Midrigan, V., and Xu, D. Y. (2018). 'How costly are markups?'. NBER Working Paper No. 24800. National Bureau of Economic Research.

Edwards, L., and van de Winkel, T. (2005). 'The market disciplining effects of trade liberalisation and regional import penetration on manufacturing in South Africa'. TIPS Working Paper No. 1-2005. TIPS.

Fedderke, J., and Hill, A. J. (2011). 'Industry structure and labor market flexibility in the South African manufacturing sector: a time series and panel data approach'. *Economic Modelling*, *28*(3): 1291–1302.

Fedderke, J., Kularatne, C., and Mariotti, M. (2007). 'Mark-up pricing in South African industry'. *Journal of African Economies*, *16*(1): 28–69.

Fedderke, J., and Naumann, D. (2011). 'An analysis of industry concentration in South African manufacturing, 1972–2001'. *Applied Economics*, *43*(22): 2919–2939.

Fedderke, J., Obikili, N., and Viegi, N. (2018). 'Markups and concentration in South African manufacturing sectors: an analysis with administrative data'. *South African Journal of Economics*, *86*(): 120–140.

Fedderke, J., and Simbanegavi, W. (2008). 'South African manufacturing industry structure and its implications for competition policy'. *Journal of Development Perspectives*, *4*(1): 134–189.

Fedderke, J., and Szalontai, G. (2009). 'Industry concentration in South African manufacturing industry: trends and consequences, 1972–96'. *Economic Modelling*, *26*(1): 241–250.

Fisher, F. M., and McGowan, J. J. (1983). 'On the misuse of accounting rates of return to infer monopoly profits'. *American Economic Review*, *73*(1): 82–97.

Fourie, F. C. v. N. (1996). 'Industrial concentration levels and trends in South Africa: completing the pictures'. *South African Journal of Economics*, *64*(1): 49–62.

Fourie, F. C. v. N., and Smit, M. R. (1989). 'Trends in economic concentration in South Africa'. *South African Journal of Economics*, *57*(3): 156–167.

Fourie, F. C. v. N., and Smith, A. (2001). 'Revisiting the concentration-profits relationship in South Africa: moving the debate beyond deadlock'. *Studies in Economics and Econometrics*, *25*(2): 25–60.

Gandhi, A., Navarro, S., and Rivers, D. (2018). 'On the identification of gross output production functions'. CHCP Working Paper No. 2018-1. Western University.

Hall, R. E. (1988). 'The relation between price and marginal cost in US industry'. *Journal of Political Economy*, *96*(5): 921–947.

Kehrig, M., and Vincent, N. (2018). 'The micro-level anatomy of the labor share decline'. NBER Working Paper No. 25275. National Bureau of Economic Research.

Kerr, A. (2018). 'Job flows, worker flows and churning in South Africa'. *South African Journal of Economics*, *86*(): 141–166.

Kreuser, C. F., and Newman, C. (2018). 'Total factor productivity in South African manufacturing firms'. *South African Journal of Economics*, *86*(): 40–78.

Leach, D. F. (1992). 'Concentration and profits in South africa: monopoly or efficiency?' *South African Journal of Economics*, *60*(2): 82–92.

Leach, D. F. (1997). 'Concentration-profits monopoly vs. efficiency debate: South African evidence'. *Contemporary Economic Policy*, *15*(2): 12–23.

Levinsohn, J., and Petrin, A. (2003). 'Estimating production functions using inputs to control for unobservables'. *The Review of Economic Studies*, *70*(2): 317–341.

Mamburu, M. (2017). 'Defining high growth firms in South Africa'. WIDER Working Paper No. 2017/107. UNU-WIDER.

Mollisi, V., and Rovigatti, G. (2017). 'Theory and practice of TFP estimation: the control function approach using Stata'. CEIS Tor Vergata Research Paper Series No. 399. University of Rome Tor Vergata.

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

Pieterse, D., Gavin, E., and Kreuser, C. F. (2018). 'Introduction to the South African Revenue Service and National Treasury firm-level panel'. *South African Journal of Economics*, *86*(): 6–39.

Ravenscraft, D. J. (1983). 'Structure-profit relationship at the line of business and industry level'. *The Review of Economics and Statistics*,(): 22–31.

Reekie, W. D. (1984). 'The structure-conduct-performance paradigm in a South African setting'. *South African Journal of Economics*, *52*(2): 97–103.

Roeger, W. (1995). 'Can imperfect competition explain the difference between primal and dual productivity measures? estimates for us manufacturing'. *Journal of Political Economy*, *103*(2): 316–330.

Rovigatti, G., and Mollisi, V. (2016, September). *PRODEST: Stata module for production function estimation based on the control function approach.* Statistical Software Components, Boston College Department of Economics. Retrieved from https://ideas.repec.org/c/boc/bocode/s458239.html

Schmalensee, R. (1989). 'Inter-industry studies of structure and performance'. In R. Schmalensee and R. D. Willig (eds), *Handbook of industrial organization* (Vol. 2, pp. 951–1009).

Shepherd, W. G. (1972). 'The elements of market structure'. *The Review of Economics and Statistics*,(): 25–37.

Shepherd, W. G., and Shepherd, J. M. (2004). *The economics of industrial organization*. Waveland Press.

Traina, J. (2018). 'Is aggregate market power increasing? production trends using financial statements'. New Working Paper No. 17. Stigler Center for the Study of the Economy and the State.

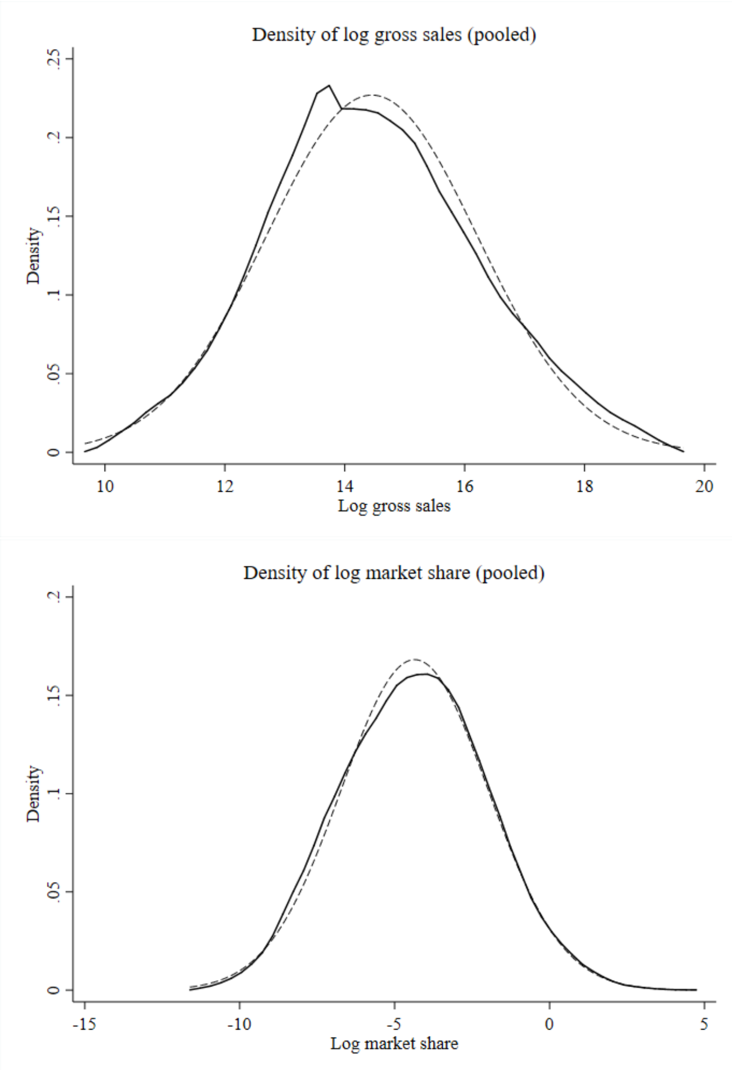
Tybout, J. R. (2000). 'Manufacturing firms in developing countries: how well do they do, and why?' *Journal of Economic Literature*, *38*(1): 11–44.

Von Broembsen, M. (2016). 'You can't bite the hand that feeds you: the commercial and contractual relations between the four large South African food retailers and their SME suppliers'. REDI3x3 Working Paper No. 22. REDI3x3.

Wier, L., and Reynolds, H. (2018). 'Big and 'unprofitable': how 10% of multinational firms do 98% of profit-shifting'. WIDER Working Paper No. 2018/111. UNU-WIDER.

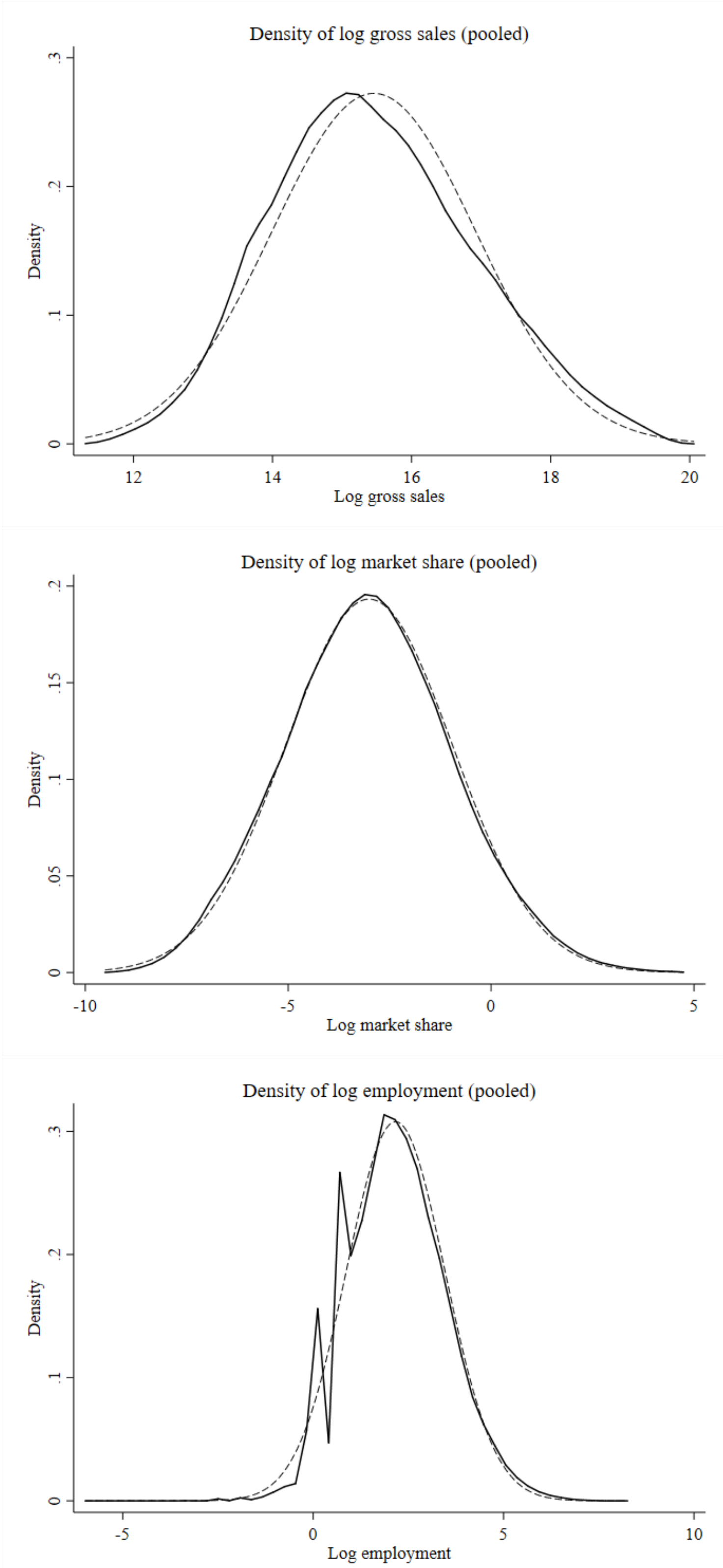
Figures

Figure 1a: Firm-size distribution, broad sample



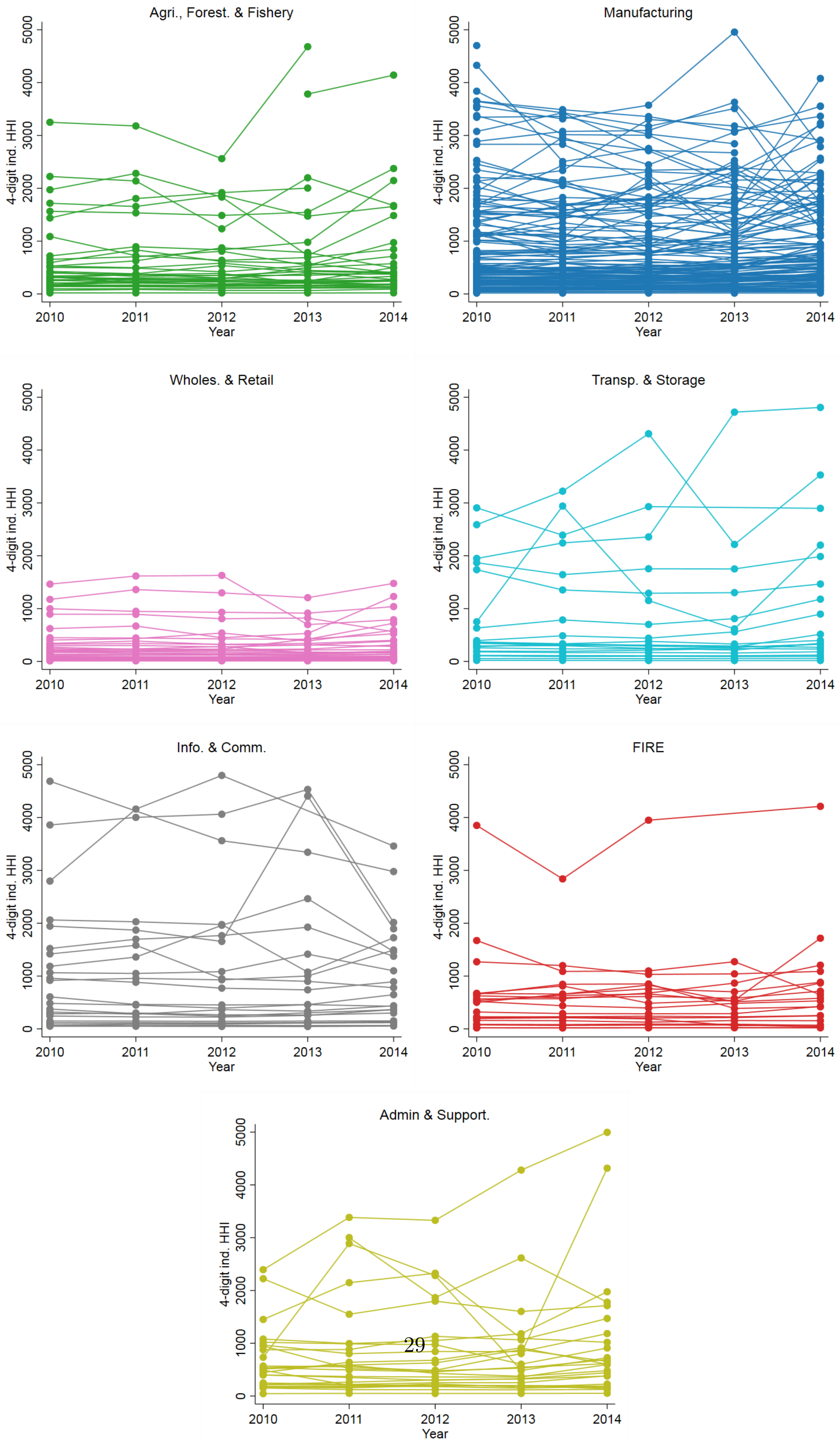
Notes: Unweighted smoothed kernel densities of firm-size variables are reported for the Broad sample. Firms are pooled across all years and sectors. The top panel shows the distribution of (log) gross sales, while the bottom panel shows the distribution of (log) market share of the 4-digit industry. Lognormal distributions are overlaid.

Figure 1b: Firm-size distribution, narrow sample



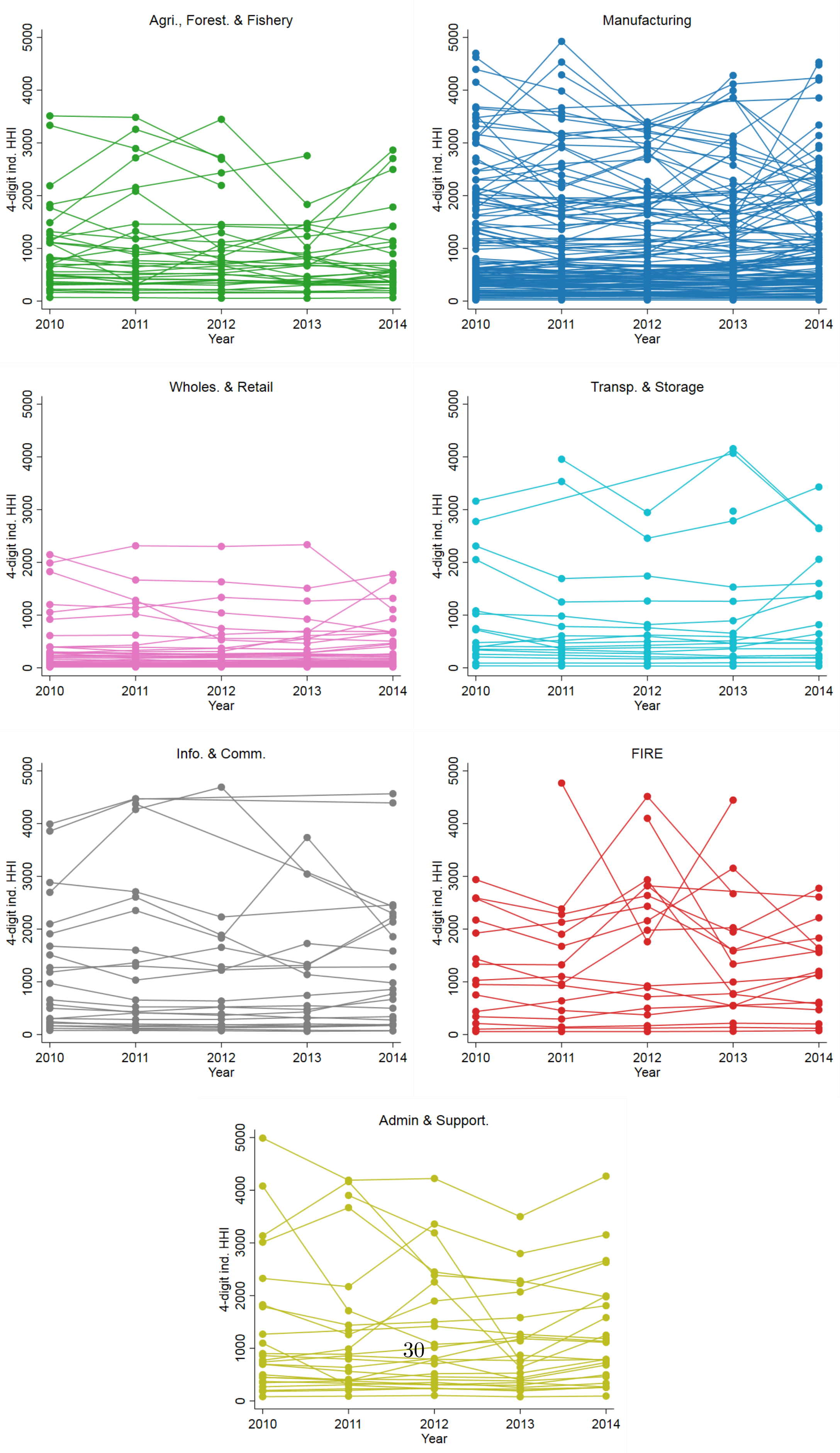
Notes: Unweighted smoothed kernel densities of firm-size variables are reported for the Narrow sample. Firms are pooled across all years and sectors. The top panel shows the distribution of (log) gross sales, the middle panel shows the distribution of (log) market share of the 4-digit industry, while the bottom panel shows the distribution of (log) full-time equivalent workers per firm. Lognormal distributions are overlaid.

Figure 2a: 4-digit HHIs over time, broad sample



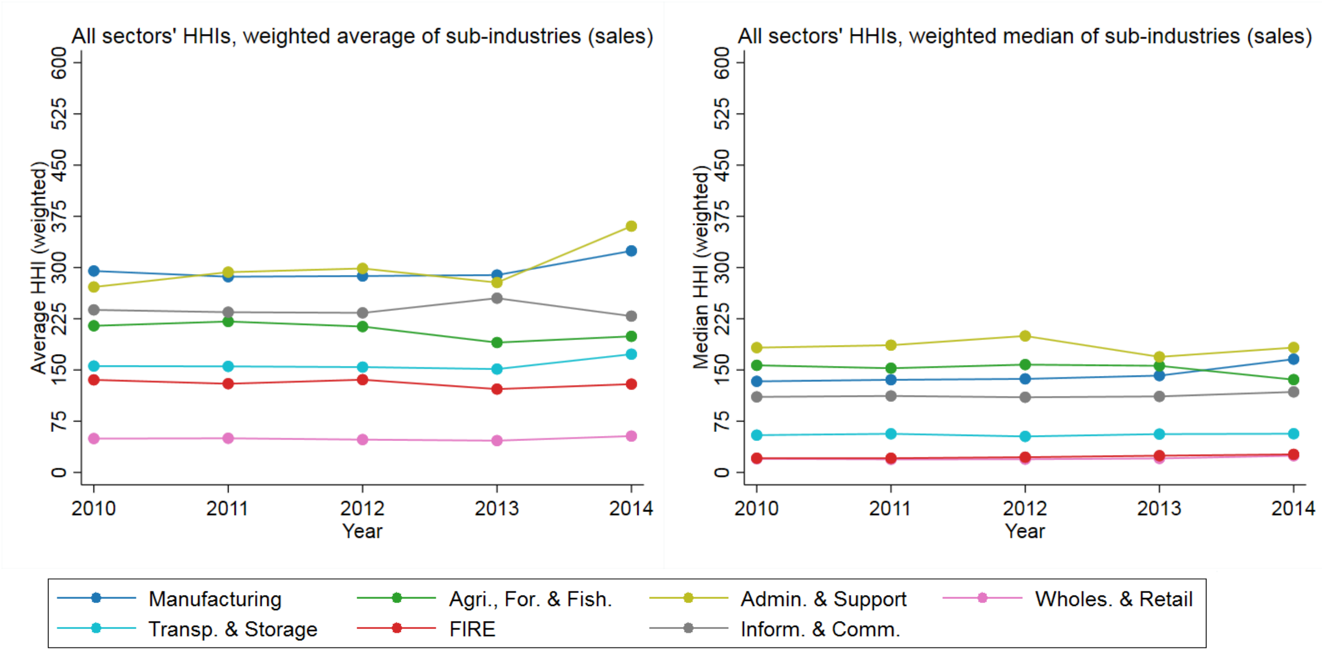
Notes: See the text.

Figure 2b: 4-digit HHIs over time, narrow sample



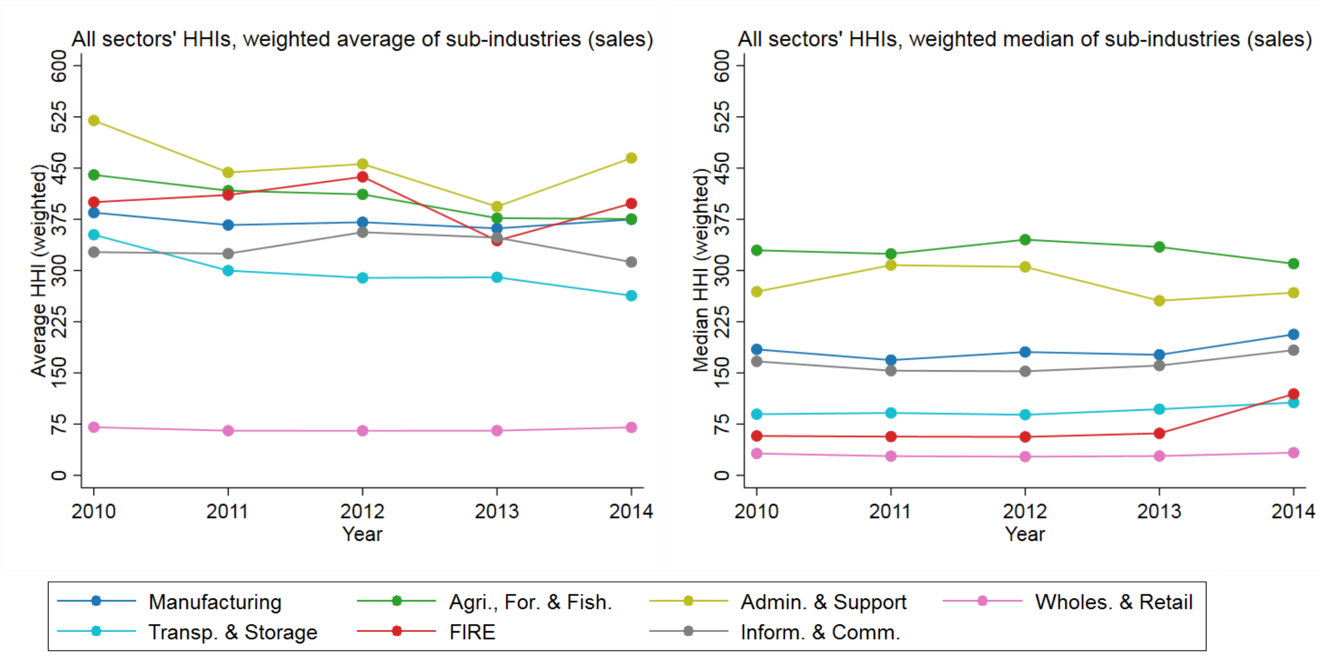
Notes: See the text.

Figure 3a: Aggregated 4-digit HHIs over time, broad sample



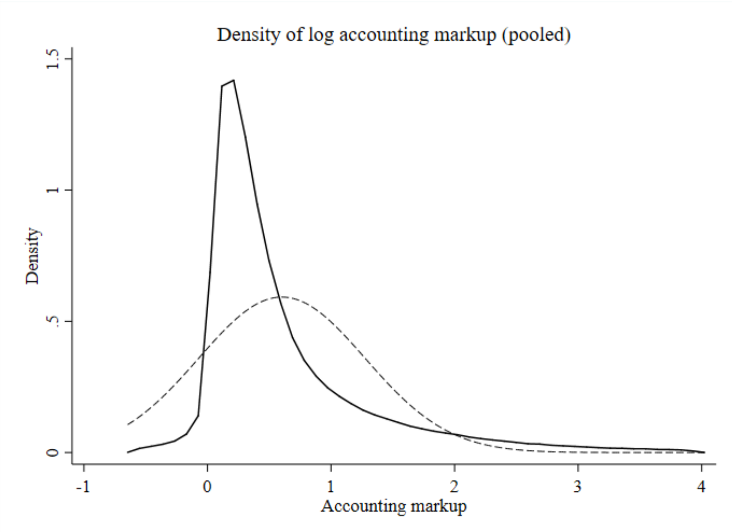
Notes: Figures show time trends in 4-digit HHI, aggregated at the 1-digit industry level, for the Broad sample. Aggregates are weighted means and medians, where weights are total gross sales per 4-digit industry per year.

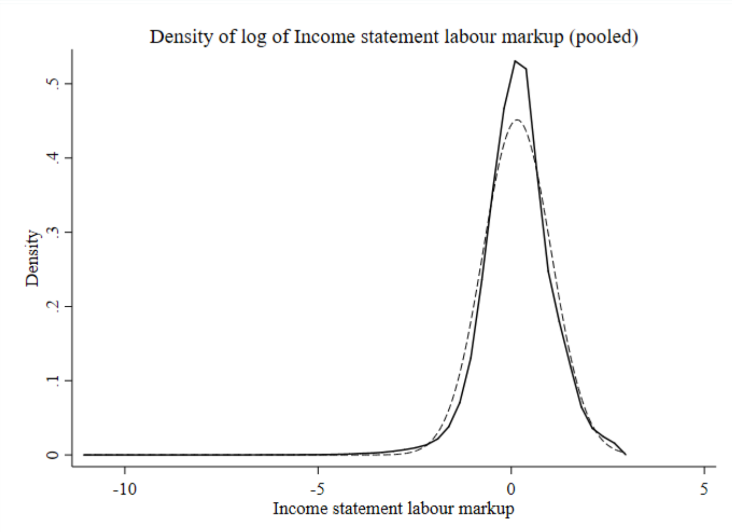
Figure 3b: Aggregated 4-digit HHIs over time, narrow sample



Notes: Figures show time trends in 4-digit HHI, aggregated at the 1-digit industry level, for the Narrow sample. Aggregates are weighted means and medians, where weights are total gross sales per 4-digit industry per year.

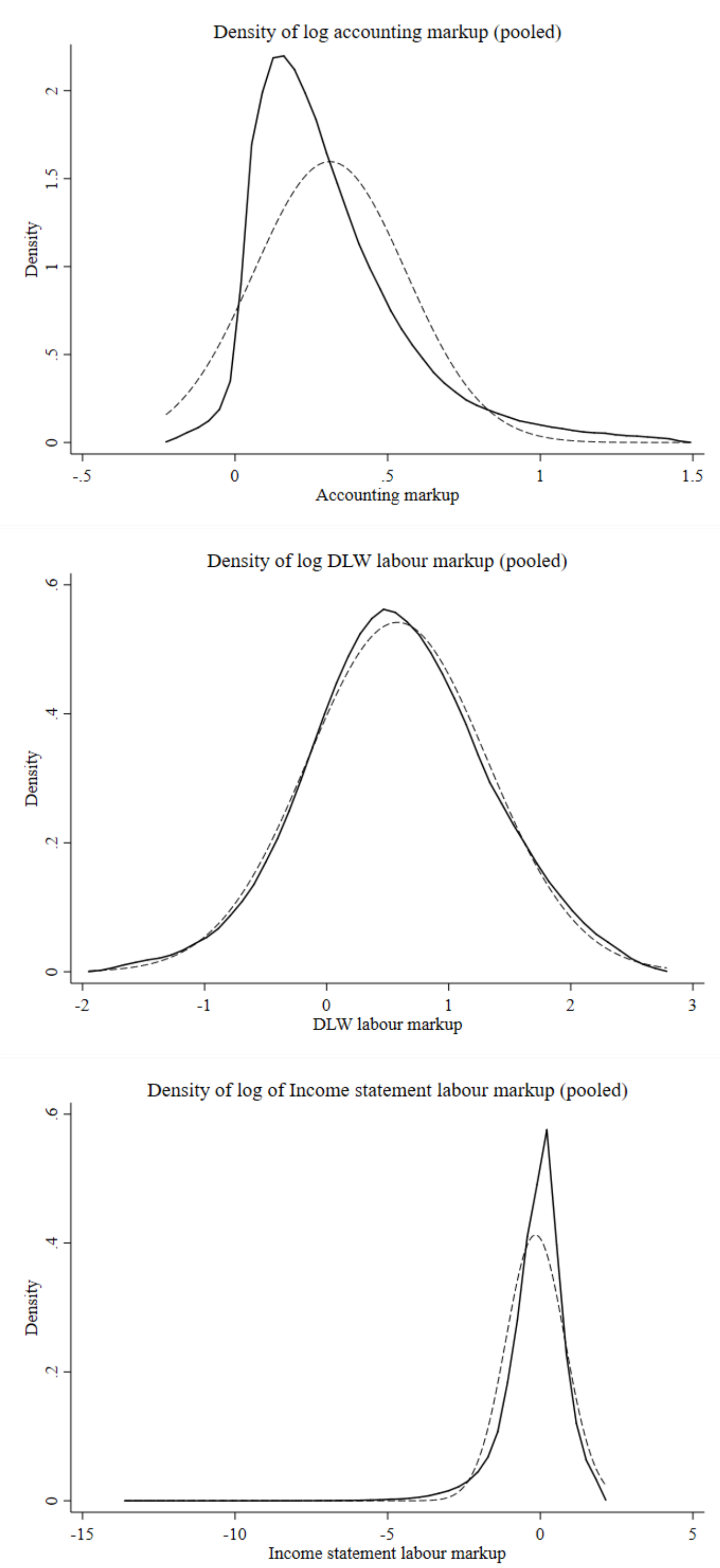
Figure 4a: Markups distribution, broad sample





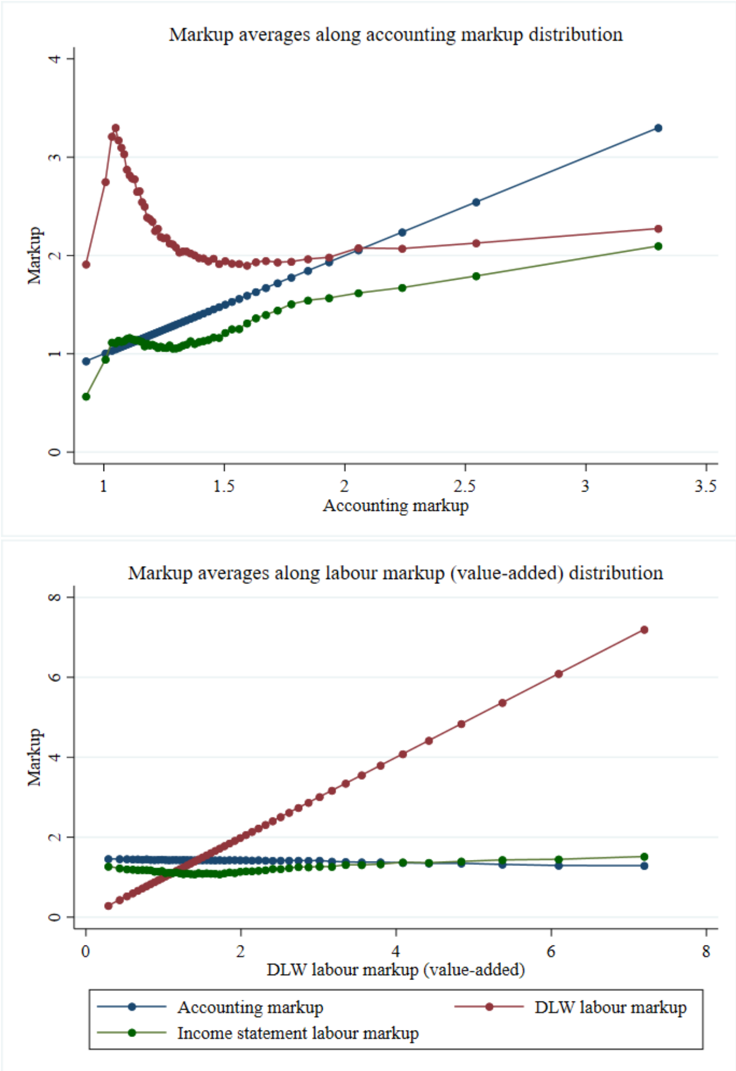
Notes: Unweighted smoothed kernel densities of different markup specifications are reported for the Broad sample. Firms are pooled across all years and sectors. The top panel shows the distribution of (log) accounting markups, while the bottom panel shows the distribution of (log) markups estimated as per De Loecker & Warzynski (2012) using income statement information on labour expenses. Lognormal distributions are overlaid.

Figure 4b: Markups distribution, narrow sample



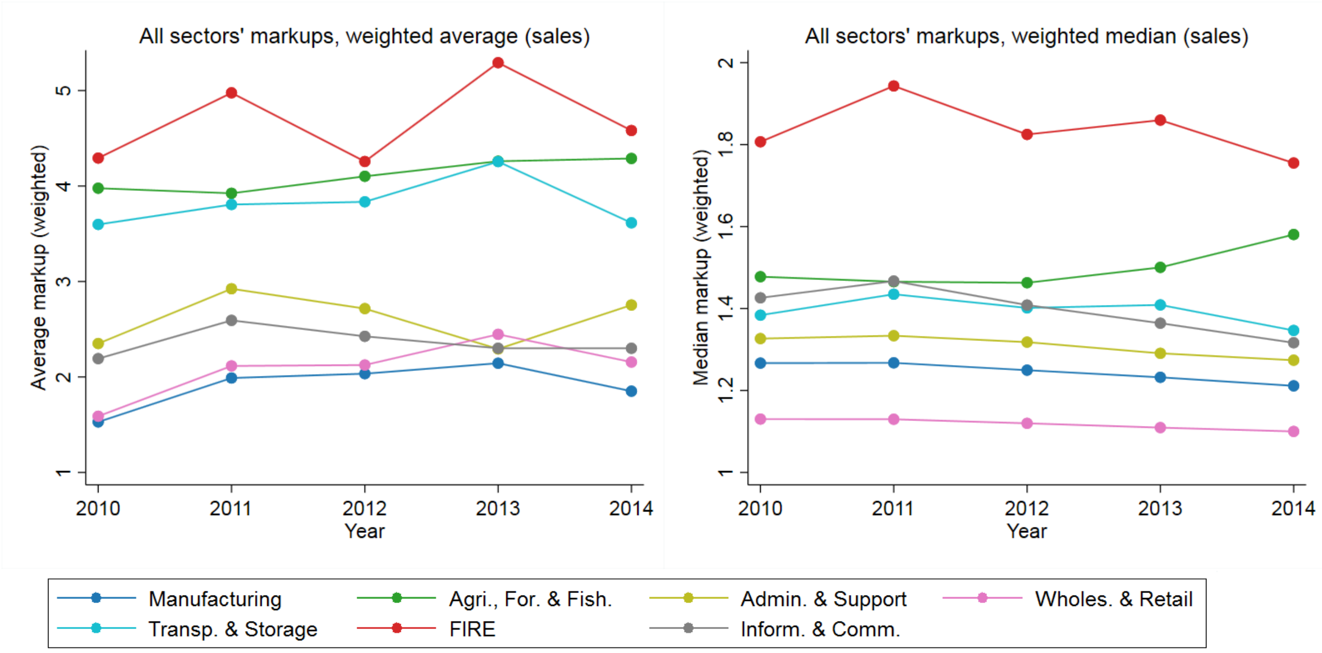
Notes: Unweighted smoothed kernel densities of different markup specifications are reported for the Narrow sample. Firms are pooled across all years and sectors. The top panel shows the distribution of (log) accounting markups, while the middle and bottom panels show the distribution of (log) markups estimated as per De Loecker & Warzynski (2012). The middle panel uses labour as the variable input, while the bottom panel uses income statement information on labour expenses. Lognormal distributions are overlaid.

Figure 5: Comparative markups, narrow sample



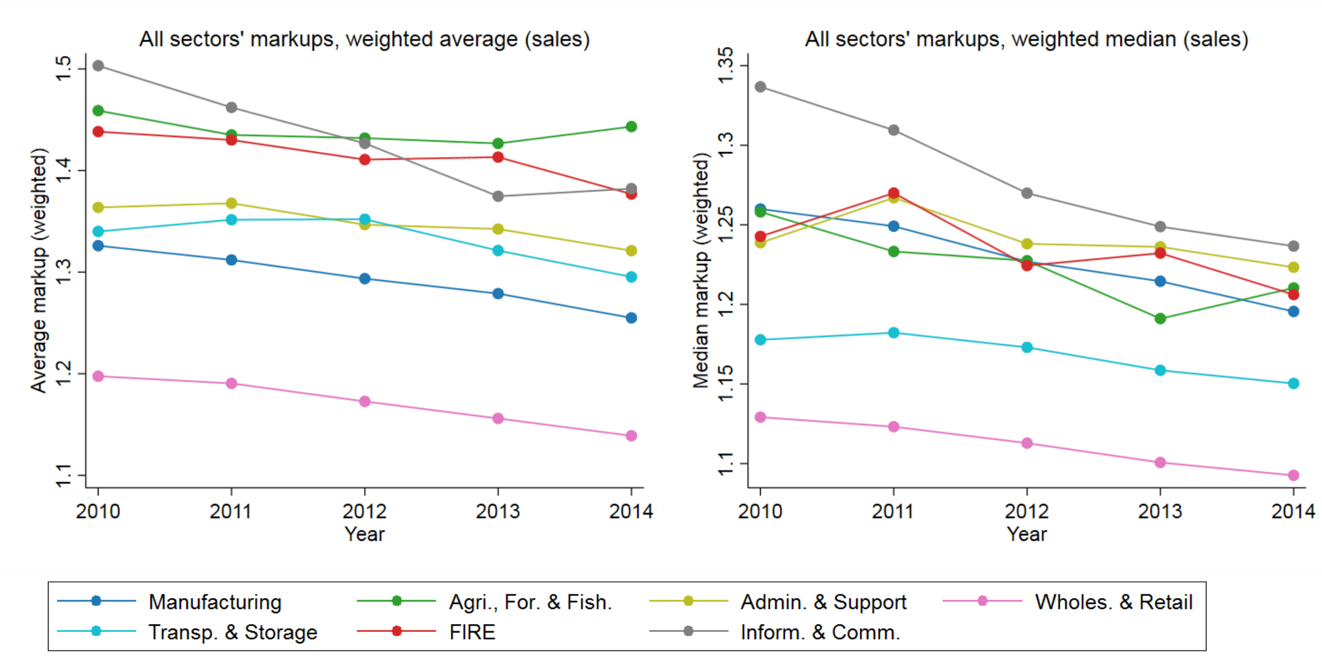
Notes: Different markup specifications are compared along 50 quantiles of 1) accounting markups and 2) estimated De Loecker & Warzynski (2012) markups using labour as the variable input. Markups are in logs.

Figure 6a: Accounting markups, broad sample



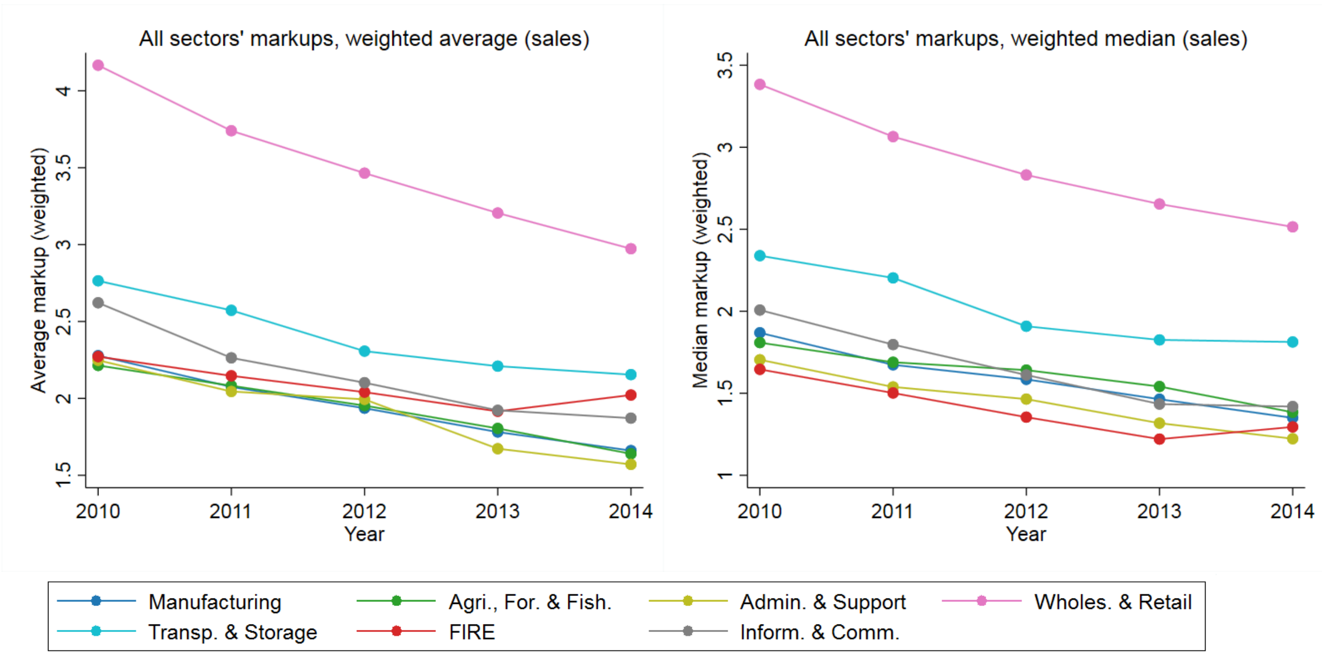
Notes: Figures show time trends in firm-level accounting markups, aggregated at the 1-digit industry level, for the Broad sample. Aggregates are weighted means and medians, where weights are gross sales per firm.

Figure 6b: Accounting markups, narrow sample



Notes: Figures show time trends in firm-level accounting markups, aggregated at the 1-digit industry level, for the Narrow sample. Aggregates are weighted means and medians, where weights are gross sales per firm.

Figure 6c: DLW labour markups, narrow sample



Notes: Figures show time trends in firm-level estimated markups, aggregated at the 1-digit industry level, for the Narrow sample. Markups are estimated as per De Loecker & Warzynski (2012), using labour as the variable input. Aggregates are weighted means and medians, where weights are gross sales per firm.

Tables

Table 1a: Firm size statistics, broad dataset

	Mean: Sales	Gini: Sales	Mean: Mkt share	Gini: Mkt share
Agri., For. & Fish	85.42	0.769	0.356	0.851
Mining & Quarry.	183.9	0.744	0.876	0.874
Manufacturing	162.1	0.731	0.682	0.850
Elec. Gas & Water	113.8	0.744	0.428	0.811
Construction	76.03	0.754	0.0564	0.830
Wholes. & Retail	149.6	0.740	0.112	0.803
Transp. & Storage	123.9	0.766	0.274	0.881
Cater. & Accom.	50.70	0.742	0.0670	0.810
Info. & Comm.	89.29	0.779	0.363	0.871
FIRE	37.71	0.812	0.0952	0.913
Prof., Sci. & Tech.	62.88	0.756	0.116	0.851
Admin & Support.	70.54	0.786	0.392	0.850
Education	33.46	0.735	0.294	0.781
Health & Soc. work	72.00	0.735	0.230	0.886
Rec. & Cult. serv.	51.32	0.810	0.319	0.889
Other services	55.32	0.780	0.101	0.948
Total	93.66	0.779	0.217	0.890

Notes: Firm size statistics are reported, aggregated across 1-digit industries with all years pooled. Mean sales are reported in 100 000s of 2014 Rands per year. Market share is of the 4-digit industry and is reported as a percentage out of 100. Results are unweighted, and are for the Broad sample of firms.

Table 1b: Firm size statistics, narrow dataset

	Mean: Sales	Gini: Sales	Mean: Mkt share	Gini: Mkt share	Mean: Empl.	Median: Empl.	Gini: Empl.
Agri., For. & Fish	176.6	0.682	1.191	0.777	40.84	13.67	0.701
Mining & Quarry.	226.5	0.669	1.574	0.828	31.40	12	0.665
Manufacturing	214.1	0.690	0.963	0.825	30.39	13.48	0.633
Elec. Gas & Water	170.4	0.692	0.739	0.792	18.32	8.250	0.623
Construction	124.0	0.693	0.118	0.770	23.09	9	0.664
Wholes. & Retail	216.9	0.688	0.183	0.768	18.12	8.167	0.620
Transp. & Storage	222.9	0.679	0.761	0.831	21.08	8.083	0.659
Cater. & Accom.	87.07	0.654	0.137	0.748	24.27	12	0.612
Info. & Comm.	130.4	0.713	0.702	0.832	15.74	6.250	0.659
FIRE	112.8	0.733	0.851	0.877	16.86	5.923	0.698
Prof., Sci. & Tech.	103.9	0.702	0.321	0.831	15.56	5.833	0.690
Admin & Support.	104.3	0.697	1.101	0.808	40.67	7.250	0.805
Education	48.77	0.651	1.055	0.726	15.40	7	0.642
Health & Soc. work	89.47	0.685	0.503	0.883	15.44	6.083	0.667
Rec. & Cult. serv.	86.65	0.728	0.867	0.854	18.35	5.134	0.737
Other services	99.17	0.710	0.244	0.938	18.46	6.500	0.700
Total	162.9	0.708	0.478	0.858	22.39	8.646	0.670

Notes: Firm size statistics are reported, aggregated across 1-digit industries with all years pooled. Mean sales are reported in 100 000s of 2014 Rands per year. Market share is of the 4-digit industry and is reported as a percentage out of 100. Employment is equivalent full-time workers per year. Results are unweighted, and are for the Narrow sample of firms.

Table 2a: HHI, broad dataset

	N sub-industries	Wgt. mean	Wgt. median	Unwgt. mean	Unwgt. median
Agri., For. & Fish	40	204.1	161.1	1529.1	372.8
Mining & Quarry.	14	407.9	126.3	2471.8	673.6
Manufacturing	137	285.0	133.3	1616.1	645.6
Elec. Gas & Water	11	205.1	76.24	817.0	682.2
Construction	12	41.90	19.95	1017.3	67.05
Wholes. & Retail	43	49.69	19.95	227.8	90.67
Transp. & Storage	21	150.1	54.59	1213.5	354.4
Cater. & Accom.	8	40.38	25.28	1387.2	90.05
Info. & Comm.	25	239.4	110.6	1872.1	882.4
FIRE	21	147.6	26.41	1167.0	514.0
Prof., Sci. & Tech.	15	88.03	30.24	1042.4	273.2
Admin & Support.	27	284.8	182.9	1047.0	534.0
Education	8	267.3	138.5	521.4	372.5
Health & Soc. work	11	99.61	37.84	2697.3	1542.5
Rec. & Cult. serv.	10	266.8	292.9	1265.0	844.6
Other services	21	38.74	4.685	1077.5	631.2
Total	424	133.9	37.84	1342.2	424.7

Notes: Summary statistics are reported for 4-digit industry HHI, aggregated across 1-digit industries with all years pooled. HHI is scaled such that if there was only one firm in the 4-digit sub-industry, the HHI would equal 10 000. The first column indicates the number of 4-digit industries within the broader 1-digit industry. Standard deviations are indicated under "SD". Where results are weighted, the weight is total aggregate sales in the 4-digit industry. Results are shown for the Broad sample.

Table 2b: HHI, narrow dataset

	N sub-industries	Wgt. mean	Wgt. median	Unwgt. mean	Unwgt. median
Agri., For. & Fish	37	425.5	305.8	1396.2	715.7
Mining & Quarry.	13	487.1	268.7	2992.7	1542.8
Manufacturing	133	364.3	168.8	1703.7	720.5
Elec. Gas & Water	11	308.7	89.43	1594.8	618.5
Construction	11	60.05	45.00	198.6	64.55
Wholes. & Retail	43	69.56	32.02	353.1	106.6
Transp. & Storage	20	295.7	89.54	1772.3	694.4
Cater. & Accom.	7	66.72	33.88	408.2	113.1
Info. & Comm.	23	321.9	153.4	1406.7	658.5
FIRE	20	352.4	57.74	1987.9	1460.2
Prof., Sci. & Tech.	15	145.6	75.24	1275.0	456.2
Admin & Support.	26	448.8	347.1	1543.7	813.2
Education	8	479.7	289.1	1046.9	776.5
Health & Soc. work	9	203.9	64.67	2076.0	1574.8
Rec. & Cult. serv.	10	399.5	103.1	1879.2	874.1
Other services	21	54.36	7.857	2138.4	826.2
Total	407	191.4	57.74	1504.9	645.0

Notes: Summary statistics are reported for 4-digit industry HHI, aggregated across 1-digit industries with all years pooled. HHI is scaled such that if there was only one firm in the 4-digit sub-industry, the HHI would equal 10 000. The first column indicates the number of 4-digit industries within the broader 1-digit industry. Standard deviations are indicated under "SD". Where results are weighted, the weight is total aggregate sales in the 4-digit industry. Results are shown for the Narrow sample.

Table 3a: Accounting markup, broad dataset

	N firms	Wgt. mean	Wgt. median	Wgt. SD	Unwgt. mean	Unwgt. median	Unwgt. SD
Agri., For. & Fish	51278	4.135	1.493	6.872	4.357	2.139	6.015
Mining & Quarry.	7501	2.664	1.328	5.380	2.599	1.448	4.028
Manufacturing	97016	1.938	1.244	4.152	1.632	1.306	2.053
Elec. Gas & Water	12696	2.039	1.132	4.539	1.933	1.333	2.955
Construction	99682	2.368	1.229	4.994	2.033	1.351	2.955
Wholes. & Retail	190458	2.132	1.116	5.104	1.524	1.211	2.083
Transp. & Storage	36809	3.854	1.397	6.615	3.833	1.986	5.057
Cater. & Accom.	58768	2.331	1.498	3.894	2.753	1.622	4.127
Info. & Comm.	32437	2.362	1.389	3.646	2.122	1.470	2.781
FIRE	99159	4.710	1.836	7.865	5.605	2.628	7.749
Prof., Sci. & Tech.	62480	2.382	1.470	3.693	2.558	1.711	3.459
Admin & Support.	32553	2.604	1.304	4.646	2.719	1.601	4.184
Education	13378	2.752	1.894	3.008	2.746	1.953	3.283
Health & Soc. work	20193	2.812	1.745	3.637	2.573	1.834	2.915
Rec. & Cult. serv.	15362	3.659	1.508	6.599	2.702	1.691	3.732
Other services	102741	2.246	1.223	4.424	2.403	1.490	3.586
Total	932511	2.488	1.215	5.123	2.650	1.451	4.267

Notes: Summary statistics are reported for firm-level accounting markups, aggregated across 1-digit industries with all years pooled. The first column indicates the number of firms within the 1-digit industry with non-missing accounting markups. Standard deviations are indicated under "SD". Where results are weighted, the weight is firm-level gross sales. Results are shown for the Broad sample.

Table 3b: Accounting markup, narrow dataset

	N firms	Wgt. mean	Wgt. median	Wgt. SD	Unwgt. mean	Unwgt. median	Unwgt. SD
Agri., For. & Fish	13652	1.438	1.220	0.580	1.650	1.406	0.693
Mining & Quarry.	3738	1.394	1.254	0.465	1.478	1.313	0.515
Manufacturing	66861	1.291	1.229	0.267	1.352	1.281	0.310
Elec. Gas & Water	7346	1.219	1.119	0.282	1.368	1.267	0.379
Construction	45798	1.288	1.187	0.353	1.391	1.284	0.402
Wholes. & Retail	116825	1.168	1.109	0.204	1.258	1.183	0.276
Transp. & Storage	12370	1.330	1.166	0.464	1.504	1.314	0.566
Cater. & Accom.	24976	1.526	1.419	0.521	1.623	1.518	0.513
Info. & Comm.	16007	1.423	1.275	0.463	1.463	1.349	0.427
FIRE	10896	1.413	1.235	0.508	1.602	1.429	0.594
Prof., Sci. & Tech.	22595	1.410	1.288	0.412	1.536	1.407	0.471
Admin & Support.	11500	1.346	1.239	0.359	1.489	1.366	0.454
Education	3700	1.666	1.530	0.502	1.730	1.614	0.533
Health & Soc. work	8740	1.560	1.412	0.530	1.687	1.567	0.561
Rec. & Cult. serv.	5574	1.493	1.291	0.546	1.619	1.444	0.581
Other services	41214	1.289	1.180	0.352	1.459	1.335	0.441
Total	411792	1.276	1.167	0.347	1.412	1.286	0.432

Notes: Summary statistics are reported for firm-level accounting markups, aggregated across 1-digit industries with all years pooled. The first column indicates the number of firms within the 1-digit industry with non-missing accounting markups. Standard deviations are indicated under "SD". Where results are weighted, the weight is firm-level gross sales. Results are shown for the Narrow sample.

Table 4: DLW markup (Labour, value-added translog), narrow dataset

	N firms	Wgt. mean	Wgt. median	Wgt. SD	Unwgt. mean	Unwgt. median	Unwgt. SD
Agri., For. & Fish	15042	1.908	1.591	1.363	1.813	1.524	1.305
Mining & Quarry.	3654	1.722	1.097	1.846	1.653	1.172	1.579
Manufacturing	66791	1.922	1.562	1.460	1.879	1.513	1.432
Elec. Gas & Water	7084	2.333	1.803	1.880	2.074	1.619	1.651
Construction	46113	2.039	1.603	1.607	2.164	1.731	1.603
Wholes. & Retail	115455	3.437	2.810	2.318	3.232	2.586	2.238
Transp. & Storage	12578	2.355	1.956	1.659	2.496	2.075	1.788
Cater. & Accom.	23823	1.386	1.122	1.062	1.278	1.041	0.955
Info. & Comm.	16040	2.119	1.602	1.767	2.336	1.782	1.892
FIRE	11396	2.071	1.384	1.979	2.320	1.648	2.095
Prof., Sci. & Tech.	22621	1.899	1.389	1.727	2.241	1.663	1.925
Admin & Support.	11157	1.873	1.423	1.600	2.039	1.515	1.832
Education	3773	1.420	1.108	1.108	1.693	1.325	1.413
Health & Soc. work	8780	2.250	1.737	1.788	2.798	2.258	1.991
Rec. & Cult. serv.	5632	1.664	1.159	1.574	1.771	1.267	1.665
Other services	41493	1.743	1.257	1.582	1.828	1.395	1.535
Total	411432	2.498	1.837	2.022	2.339	1.762	1.905

Notes: Summary statistics are reported for estimated firm-level markups, aggregated across 1-digit industries with all years pooled. Markups are estimated as per De Loecker & Warzynski (2012), using labour as the variable input. The first column indicates the number of firms within the 1-digit industry with non-missing markups. Standard deviations are indicated under “SD”. Where results are weighted, the weight is firm-level gross sales. Results are shown for the Narrow sample.

Table 5a: Income statement markup (Labour, value-added translog), broad dataset

	N firms	Wgt. mean	Wgt. median	Wgt. SD	Unwgt. mean	Unwgt. median	Unwgt. SD
Agri., For. & Fish	27068	4.430	3.557	3.238	3.860	2.871	3.357
Mining & Quarry.	4675	1.566	1.171	1.537	1.135	0.823	1.239
Manufacturing	81235	1.747	1.255	1.567	1.400	0.947	1.503
Elec. Gas & Water	10312	1.962	1.632	1.561	1.671	1.336	1.615
Construction	74086	1.824	1.156	1.938	1.445	0.889	1.707
Wholes. & Retail	167223	1.707	1.498	1.045	1.462	1.230	1.161
Transp. & Storage	13251	1.022	0.857	1.252	1.097	0.665	1.607
Cater. & Accom.	43522	3.779	3.781	2.069	2.992	2.782	2.174
Info. & Comm.	22159	1.136	0.803	1.252	0.866	0.536	1.197
FIRE	18992	1.515	1.094	1.618	1.338	0.827	1.791
Prof., Sci. & Tech.	31985	1.622	1.309	1.461	1.478	1.028	1.718
Admin & Support.	17289	1.650	1.102	1.591	1.298	0.778	1.609
Education	6405	0.782	0.565	1.079	0.703	0.435	1.088
Health & Soc. work	12791	1.148	0.895	0.936	1.089	0.834	1.047
Rec. & Cult. serv.	10196	1.855	1.335	1.636	1.456	0.957	1.767
Other services	65629	2.308	1.971	1.549	2.123	1.638	1.860
Total	606818	1.893	1.459	1.678	1.687	1.165	1.810

Notes: Summary statistics are reported for estimated firm-level markups, aggregated across 1-digit industries with all years pooled. Markups are estimated as per De Loecker & Warzynski (2012), using labour expenses from the firm income statement as the variable input. The first column indicates the number of firms within the 1-digit industry with non-missing markups. Standard deviations are indicated under "SD". Where results are weighted, the weight is firm-level gross sales. Results are shown for the Broad sample.

Table 5b: Income statement markup (Labour, gross output translog), narrow dataset

	N firms	Wgt. mean	Wgt. median	Wgt. SD	Unwgt. mean	Unwgt. median	Unwgt. SD
Agri., For. & Fish	14220	1.150	0.784	1.064	0.989	0.669	1.013
Mining & Quarry.	3335	1.613	1.394	0.945	1.165	0.942	0.842
Manufacturing	56133	1.536	1.121	1.342	1.117	0.710	1.141
Elec. Gas & Water	7110	1.621	1.402	0.862	1.261	1.130	0.763
Construction	42673	1.567	1.249	1.045	1.210	0.948	0.909
Wholes. & Retail	89443	0.987	0.966	0.669	0.920	0.870	0.603
Transp. & Storage	11622	2.348	1.960	1.201	2.155	1.784	1.327
Cater. & Accom.	24635	2.146	2.094	0.983	1.767	1.700	0.995
Info. & Comm.	15621	1.385	1.108	1.015	1.046	0.853	0.779
FIRE	9658	0.824	0.374	1.002	0.788	0.546	0.910
Prof., Sci. & Tech.	19574	1.047	0.685	0.960	0.847	0.483	0.992
Admin & Support.	9085	1.190	0.955	1.261	1.023	0.784	1.114
Education	3582	1.977	1.679	1.134	1.719	1.393	1.160
Health & Soc. work	7694	3.690	3.412	1.374	3.627	3.439	1.602
Rec. & Cult. serv.	5364	1.114	0.692	1.022	0.490	0.390	0.439
Other services	39096	1.694	1.545	0.787	1.454	1.261	0.893
Total	358845	1.381	1.206	1.098	1.214	0.964	1.048

Notes: Summary statistics are reported for estimated firm-level markups, aggregated across 1-digit industries with all years pooled. Markups are estimated as per De Loecker & Warzynski (2012), using labour expenses from the firm income statement as the variable input. The first column indicates the number of firms within the 1-digit industry with non-missing markups. Standard deviations are indicated under "SD". Where results are weighted, the weight is firm-level gross sales. Results are shown for the Broad sample.

Table 6: Regression results: Accounting markups

	Broad sample							Narrow sample						
	Agri.	Manuf.	W. & Ret.	Tr. & Sto.	ICT	FIRE	Admin.	Agri.	Manuf.	W. & Ret.	Tr. & Sto.	ICT	FIRE	Admin.
Coefficients														
Concentration	−0.164***	−0.007	0.019***	−0.023	−0.001	−0.059**	0.012	−0.013	−0.023***	0.003	0.013**	−0.017***	0.002	0.010
	(0.060)	(0.006)	(0.005)	(0.021)	(0.014)	(0.023)	(0.022)	(0.011)	(0.004)	(0.003)	(0.006)	(0.006)	(0.009)	(0.009)
Market share	0.233***	0.004	−0.022***	0.033*	0.009	0.087***	−0.024	0.012	0.026***	−0.004	−0.021**	−0.006	−0.012	0.011
	(0.071)	(0.009)	(0.004)	(0.019)	(0.028)	(0.019)	(0.031)	(0.023)	(0.005)	(0.004)	(0.007)	(0.018)	(0.028)	(0.023)
MshareXConc.	−0.016	0.001	0.001	0.000	0.002	−0.005	0.005	0.003	−0.001	−0.001	0.001	0.006	0.004	−0.003
	(0.011)	(0.002)	(0.001)	(0.004)	(0.005)	(0.004)	(0.006)	(0.004)	(0.001)	(0.001)	(0.002)	(0.003)	(0.005)	(0.004)
Marginal effects														
Concentration														
Median mshare	−0.116***	−0.009	0.014***	−0.024	−0.009	−0.032**	−0.002	−0.017*	−0.022***	0.006*	0.012*	−0.030***	−0.007	0.015*
	(0.029)	(0.006)	(0.005)	(0.016)	(0.015)	(0.015)	(0.013)	(0.010)	(0.004)	(0.003)	(0.007)	(0.010)	(0.012)	(0.008)
Market share														
Median conc.	0.154***	0.007	−0.018***	0.034***	0.020*	0.071***	0.002	0.030***	0.023***	−0.008***	−0.018***	0.023***	0.007	−0.009*
	(0.022)	(0.005)	(0.003)	(0.009)	(0.011)	(0.011)	(0.010)	(0.008)	(0.003)	(0.003)	(0.005)	(0.007)	(0.007)	(0.006)
Low conc.	0.187***	0.005	−0.019***	0.034***	0.018	0.072***	−0.005	0.025**	0.024***	−0.007**	−0.018***	0.018***	0.004	−0.005
	(0.041)	(0.006)	(0.003)	(0.010)	(0.013)	(0.011)	(0.011)	(0.011)	(0.004)	(0.003)	(0.004)	(0.007)	(0.009)	(0.005)
High conc.	0.140***	0.007	−0.016***	0.034***	0.023*	0.059***	0.007	0.034***	0.022***	−0.009***	−0.017***	0.029***	0.015	−0.013
	(0.017)	(0.006)	(0.005)	(0.013)	(0.013)	(0.010)	(0.015)	(0.009)	(0.003)	(0.003)	(0.006)	(0.008)	(0.010)	(0.009)
N	47664	91788	180127	33825	29817	87247	29087	12352	63762	111066	11322	14530	9144	10069
Clusters	39	136	43	21	24	21	26	37	132	43	20	23	20	26
Adj. R^2	0.629	0.533	0.556	0.735	0.583	0.740	0.685	0.687	0.643	0.728	0.766	0.634	0.647	0.705
Adj. Within R^2	0.029	0.000	0.001	0.003	0.001	0.005	−0.000	0.005	0.005	0.001	0.011	0.005	0.000	0.001

Notes: Firm-level fixed effect regressions of logged accounting markups on logged regressors are shown. Each column is a seperate regression estimated for the given industry. Marginal effects at given values of covariates are indicated in the bottom half of the table. "High" concentration means a marginal effect of market share at the 90th percentile of concentration in the estimation sample, while "Low" concentration is the marginal effect at the comparable 10th percentile. Sectors correspond to those discussed in the text. The left super-column shows estimates for the broad sample while the right super-column shows estimates for the narrow sample. Control variables are not shown.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 7: Regression results: DLW markups

	Agri.	Manuf.	W. & Ret.	Tr. & Sto.	ICT	FIRE	Admin.
Coefficients							
Concentration	−0.023	−0.059***	−0.017	−0.137**	−0.046	0.050	0.023
	(0.023)	(0.016)	(0.024)	(0.050)	(0.040)	(0.047)	(0.029)
Market share	0.064*	0.167***	−0.011	0.002	0.003	−0.058	−0.109
	(0.033)	(0.026)	(0.017)	(0.058)	(0.040)	(0.049)	(0.070)
MshareXConc.	0.002	−0.011**	0.026***	0.023**	0.006	0.015	0.015
	(0.005)	(0.005)	(0.004)	(0.009)	(0.009)	(0.009)	(0.011)
Marginal effects							
Concentration							
Median mshare	−0.026	−0.038***	−0.099***	−0.188***	−0.059	0.013	−0.002
	(0.021)	(0.015)	(0.027)	(0.047)	(0.039)	(0.048)	(0.032)
Market share							
Median conc.	0.076***	0.112***	0.083***	0.106***	0.033*	0.014	−0.026
	(0.015)	(0.010)	(0.011)	(0.035)	(0.018)	(0.016)	(0.034)
Low conc.	0.072***	0.133***	0.062***	0.084**	0.028*	0.003	−0.044
	(0.018)	(0.014)	(0.011)	(0.038)	(0.017)	(0.018)	(0.036)
High conc.	0.078***	0.096***	0.121***	0.146***	0.040*	0.044*	−0.010
	(0.015)	(0.011)	(0.013)	(0.035)	(0.024)	(0.024)	(0.036)
N	13781	63708	109625	11485	14557	9550	9749
Clusters	37	131	43	20	23	20	26
Adj. R^2	0.748	0.810	0.841	0.786	0.785	0.809	0.829
Adj. Within R^2	0.064	0.035	0.022	0.033	0.014	0.045	0.003

Notes: Firm-level fixed effect regressions of logged estimated markups on logged regressors are shown. Markups are estimated as per De Loecker & Warzynski (2012), using labour as the variable input. Each column is a seperate regression estimated for the given industry. Marginal effects at given values of covariates are indicated in the bottom half of the table. "High" concentration means a marginal effect of market share at the 90th percentile of concentration in the estimation sample, while "Low" concentration is the marginal effect at the comparable 10th percentile. Sectors correspond to those discussed in the text. Control variables are not shown.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table 8: Regression results: DLW markups with ω productivity

	Agri.	Manuf.	W. & Ret.	Tr. & Sto.	ICT	FIRE	Admin.
Coefficients							
Concentration	0.018	0.010	−0.020	−0.136**	−0.001	0.153*	0.048*
	(0.019)	(0.031)	(0.026)	(0.063)	(0.045)	(0.075)	(0.027)
Market share	−0.004	0.161***	−0.010	−0.032	−0.029	−0.177**	−0.159***
	(0.036)	(0.028)	(0.017)	(0.072)	(0.053)	(0.071)	(0.053)
MshareXConc.	−0.001	−0.022***	0.026***	0.025**	−0.012	0.016	0.007
	(0.005)	(0.004)	(0.004)	(0.010)	(0.010)	(0.009)	(0.010)
ω productivity	0.667***	0.302***	0.061**	0.167***	0.743***	0.787***	0.600***
	(0.079)	(0.051)	(0.025)	(0.035)	(0.073)	(0.050)	(0.113)
Marginal effects							
Concentration							
Median mshare	0.019	0.053**	−0.101***	−0.192***	0.026	0.114**	0.036
	(0.017)	(0.027)	(0.029)	(0.061)	(0.042)	(0.056)	(0.034)
Market share							
Median conc.	−0.010	0.052***	0.082***	0.082*	−0.092***	−0.101***	−0.118***
	(0.017)	(0.013)	(0.012)	(0.046)	(0.028)	(0.032)	(0.040)
Low conc.	−0.008	0.094***	0.061***	0.058	−0.082***	−0.113***	−0.126***
	(0.020)	(0.018)	(0.012)	(0.049)	(0.027)	(0.037)	(0.036)
High conc.	−0.011	0.018	0.121***	0.126***	−0.107***	−0.070***	−0.110**
	(0.018)	(0.011)	(0.015)	(0.045)	(0.034)	(0.023)	(0.045)
N	13482	62087	106471	11194	14257	9442	9667
Clusters	37	131	43	20	23	20	26
Adj. R^2	0.760	0.819	0.839	0.791	0.806	0.833	0.842
Adj. Within R^2	0.105	0.068	0.019	0.039	0.118	0.125	0.078

Notes: See Table 7 for the baseline specification. In the regression reported here, a structurally derived productivity proxy, ω, is included as an additional control.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Appendix A: Markup estimation per De Loecker and Warzynski (2012)

An expression for markups.—A firm f produce output Q_{ft} in time t according to the production technology $Q_{ft}(\cdot)$ such that

$$Q_{ft} = Q_{ft}(X_{ft}^1, \ldots, X_{ft}^V, K_{ft}, \omega_{ft}), \tag{6}$$

where in the production process the firm uses V variable inputs X^v (such as labour and materials) in addition to the capital stock K_{ft}, which is taken to be a dynamic input into production. ω_{ft} is firm-specific Hicks-neutral unobserved productivity. Assuming producers are cost-minimising, the associated Lagrangian function $L(\cdot)$ is

$$L(X_{ft}^1, \ldots, X_{ft}^V, K_{ft}, \lambda_{ft}) = \sum_{v=1}^{V} P_{ft}^{X^v} X_{ft}^v + r_{ft} K_{ft} + \lambda_{ft}(Q_{ft} - Q_{ft}(\cdot)) \tag{7}$$

where $P_{ft}^{X^v}$ and r_{ft} are the firm's input prices for variable input X^v and capital respectively. The first-order condition for any of the variable inputs (assumed to be free of adjustment costs) is

$$\frac{\partial L_{ft}}{\partial X_{ft}^v} = P_{ft}^{X^v} - \lambda_{ft} \frac{\partial Q_{ft}(\cdot)}{\partial X_{ft}^v} = 0. \tag{8}$$

The Lagrangian formulation makes it apparent that λ_{ft} is the marginal cost of production at a given level of output, as $\frac{\partial L_{ft}}{\partial Q_{ft}} = \lambda_{ft}$. Defining the markup as the price-cost ratio $\frac{P}{MC}$, so that $\mu_{ft} = \frac{P}{MC} = \frac{P_{ft}}{\lambda_{ft}}$, rearranging equation 8 and multiplying both sides by X_{ft}^v yields equation 3. As discussed above, with the intution of these results, this can then be re-arranged to equation 4.

Estimation.—A preliminary issue concerns the form of the production function, which needs to be explicitly specified. A first choice concerns whether a gross output or value-added specification is used. For my main specification I use a value-added function, given concerns about the identification of gross output production function parameters in an ACF setting (Ackerberg et al. 2015; Gandhi et al. 2018).[37] A second choice concerns functional form. The most common choices are the standard Cobb-Douglas production function and the more flexible translog production function.[38] A disadvantage of the Cobb-Douglas specification is that output elasticities are constrained to be constant across the levels of inputs used. This is particularly disadvantageous for the analysis at hand, as it means that for each sample over which the production function is estimated, firm-level markups will only differ by their individual input revenue shares. I therefore instead specify translog production functions.

In their discussion of estimation, De Loecker and Warzynski (2012) start from a more general production function

$$Q_{ft} = G(X_{ft}^1, \ldots, X_{ft}^V, K_{ft}; \beta) \exp(\omega_{ft}), \tag{9}$$

whereas below I quickly move to the parametric specification which I implement. However it is useful to use their expression to highlight a few features of their model set-up. Firstly, their setup highlights that I estimate a set of technology parameters β which are common to all firms over which markups are

[37]A value-added output is equivalent to gross output minus material inputs, and allows specification of a production function where materials do not enter as inputs. Under a Cobb-Douglas specification, a value added production function would be $q_{ft} = \beta_0 + \beta_k k_{ft} + \beta_l l_{ft} + \omega_{ft} + \varepsilon_{ft}$ whereas a gross output function would specify a material input such that $q_{ft} = \beta_0 + \beta_k k_{ft} + \beta_l l_{ft} + \beta_m m_{ft} + \omega_{ft} + \varepsilon_{ft}$. The terms q k, l and m are respectively the logarithms of output, the capital stock, labour, and materials. ACF discuss how the existence of a meaningful value-added production function is not obvious in all circumstances. While my main specification uses a value-added production function, I also estimate gross output specifications and compare results.

[38]A value-added Cobb-Douglas production function is outlined in footnote 37, while the value-added translog specification is given in equation 5.

estimated. Secondly, I explicitly allow for measurement error in observed output and for unanticipated shocks to production, which are combined into ε_{ft}. Observed logged output is given by $y_{ft} = \ln(Q_{ft}) + \varepsilon_{ft}$. This is an important point to note, as measurement error will be purged from output as part of the estimation procedure.

As discussed, in my main value-added translog specification I estimate a production function of the form of equation 5. The chief concern preventing straightforward OLS estimations has been the simultaneity bias caused by absorbing unobserved productivity ω_{ft} into a composite error term $u_{ft} = \omega_{ft} + \varepsilon_{ft}$. While the econometrician does not observe ω_{ft}, the firm does, and its input choices will be affected by its productivity and therefore endogenous.

The currently most-popular approach for production function estimation, pioneered by Olley and Pakes (1996), is to use a two-step control function approach. The ACF approach outlined here is a refinement of the Olley and Pakes (1996) method. The firm's demand for material inputs is taken to be of the form

$$m_{ft} = m_t(k_{ft}, l_{ft}, \omega_{ft}, \mathbf{z}_{ft}), \tag{10}$$

where additional control variables affecting the firm's input demand are captured in the vector $\mathbf{z}_{ft}$.[39] The key assumption is that $m_t(k_{ft}, l_{ft}, \omega_{ft}, \mathbf{z}_{ft})$ is strictly increasing in ω_{ft}. This allows an inversion of $m_t(\cdot)$ to get

$$\omega_{ft} = h_t(k_{ft}, l_{ft}, m_{ft}, \mathbf{z}_{ft}), \tag{11}$$

where $h_t(\cdot)$ will serve as a proxy for productivity.

In the first stage, in order to obtain predicted output $\hat{\varphi}_{ft}$ and an estimate for the error ε_{ft}, I run the regression

$$y_{ft} = \varphi_t(k_{ft}, l_{ft}, m_{ft}, \mathbf{z}_{ft}) + \varepsilon_{ft}, \tag{12}$$

where as per equations 5 and 11

$$\varphi_t(k_{ft}, l_{ft}, m_{ft}, \mathbf{z}_{ft}) = \beta_l l_{ft} + \beta_k k_{ft} + \beta_{ll} l_{ft}^2 + \beta_{kk} k_{ft}^2 + \beta_{lk} l_{ft} k_{ft} + h_t(k_{ft}, l_{ft}, m_{ft}, \mathbf{z}_{ft}) \tag{13}$$

and the non-parametric function h_t is approximated by a high-degree polynomial in its arguments.

With the quantity $\hat{\varphi}_{ft}$ in hand, I then specify a law of motion for productivity growth and (following De Loecker et al. (2018)) assume that it follows an AR(1) process, such that $\omega_{ft} = \rho\omega_{f,t-1} + \xi_{ft}$. From equations 11 and 13, which together give an expression for ω_{ft} in terms of φ_{ft} and the translog inputs l_{ft}, k_{ft}, l_{ft}^2, k_{ft}^2, and $l_{ft}k_{ft}$, I can then project ω_{ft} on it's lag such that

$$\begin{aligned} &\hat{\varphi}_{ft} - \beta_l l_{ft} - \beta_k k_{ft} - \beta_{ll} l_{ft}^2 - \beta_{kk} k_{ft}^2 - \beta_{lk} l_{ft} k_{ft} = \\ &\rho(\hat{\varphi}_{f,t-1} - \beta_l l_{f,t-1} - \beta_k k_{f,t-1} - \beta_{ll} l_{f,t-1}^2 - \beta_{kk} k_{f,t-1}^2 - \beta_{lk} l_{f,t-1} k_{f,t-1}) + \xi_{ft}. \end{aligned} \tag{14}$$

From equation 14, I can now form moments to obtain estimates of the production function parameters with standard GMM techniques. Specifically I rely on

$$\mathbb{E}\left(\xi_{ft}(\beta)\begin{pmatrix} l_{f,t-1} \\ k_{ft} \\ l_{f,t-1}^2 \\ k_{ft}^2 \\ l_{f,t-1}k_{ft} \end{pmatrix}\right) = 0. \tag{15}$$

These moment conditions are valid under the assumption that capital is decided a period ahead and is therefore not correlated with the contemporaneous productivity shock, while lagged labour is used to

[39] In $\mathbf{z}_{ft}$ I include firm-level 4-digit industry market share, firm age, and log of the capital/sales ratio.

identify the labour coefficients because labour is assumed to freely adjustable and therefore reactive to shocks. In order for lagged labour to be a valid instrument for current labour, input prices must be correlated over time, for which De Loecker and Warzynski (2012) find very strong evidence.

The above provides an estimate of the output elasticity of labour, overcoming the main obstacle to markup estimation as per 4. However rather than calculating the expenditure share α^X_{ft} directly from the data, I use the estimate of ε_{ft} obtained from the first stage regression (equation 13) to purge the observed output of measurement error and other variation in expenditure shares not related to input demand. As such I calculate the expenditure share according to $\alpha^X_{ft} = (P^X_{ft} X_{ft}) / \left(P_{ft} \frac{\widetilde{Q}_{ft}}{\exp(\hat{\varepsilon}_{ft})} \right)$, where $\widetilde{Q}_{ft}$ is the output observed in the data.

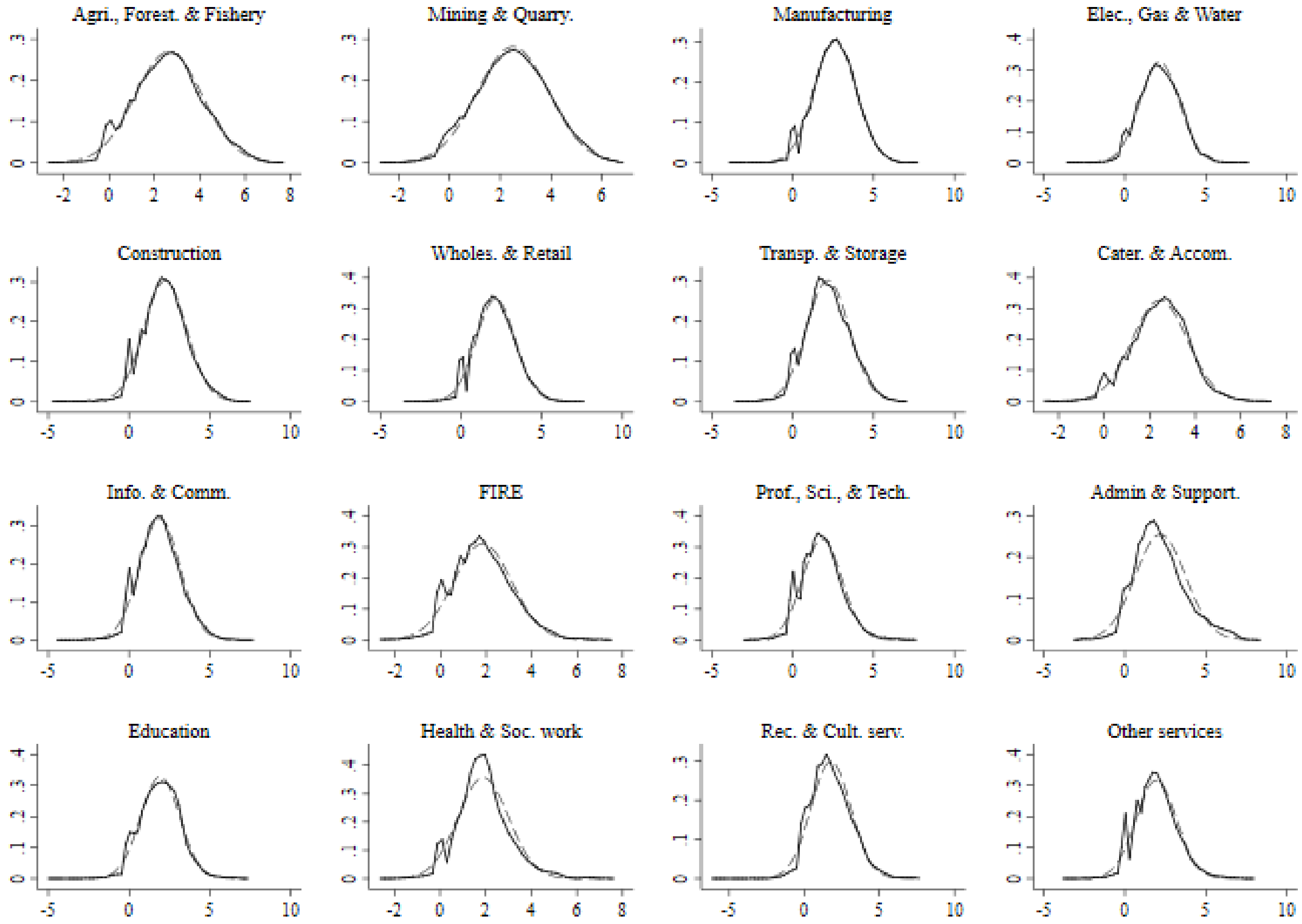
Lastly, the approach is amenable to an estimate for firm-specific productivity, as discussed above. As per equations 11, 13 and 14, $\hat{\omega}_{ft} = \hat{\varphi}_{ft} - \hat{\beta}_l l_{ft} - \hat{\beta}_k k_{ft} - \hat{\beta}_{ll} l^2_{ft} - \hat{\beta}_{kk} k^2_{ft} - \hat{\beta}_{lk} l_{ft} k_{ft}$ where the estimated coefficients β come from the second-stage, equation 14.

Appendix B: Other markup specifications

Apart from my main DLW and Income statement markups described in Section 4.3, I estimate a few additional markup specifications. Firstly I implement the DLW and Income statement approaches above with gross output specifications, and estimate markups over labour and materials. Secondly, I attempt also to estimate markups as per De Loecker et al. (2018), who use a bundled variable costs variable as both the variable input *and* the materials proxy. When implementing this approach I use cost of sales plus labour expenses as this input. These additional, unreported specifications exhibit significant volatility in the estimated markups. Standard deviations are frequently larger than means, sometimes dramatically so, and it is often the case that between 10 and 50% of markups are calculated as *negative* according to these specifications. The only exceptions are gross output income statement approaches estimated over the broad sample, where *more* than 50% of markups are found to be negative. Negative markups indicate a serious problem with the method or data. They are nonsensical economically, and come from negative output elasticities estimated in the production function, which would seem to violate the necessary assumption for markup estimation of cost-minimizing firms. Given the theoretical and practical problems with these specifications, I do not report their results, and stick to my value-added DLW and Income statement approaches.For my preferred DLW value-added approach, with markups estimated over labour, fewer than 1% of observations are negative and are dropped as part of the routine trimming procedure. For the value-added Income statment approaches, less than 5% of markups are negative. All negative markups are by necessity dropped from the dataset.

Appendix C

Figure C1: Industry-specific densities of log employment, narrow sample



Notes: Unweighted smoothed kernel densities of log employment are reported for each 1-digit industry in Narrow sample. Firms are pooled across all years. Specifically this is the distribution of (log) full-time equivalent workers per firm. Lognormal distributions are overlaid.