Journal of Productivity Analysis (2020) 53:243–263
https://doi.org/10.1007/s11123-019-00571-8

Non-structural and structural models in productivity analysis: study of the British Isles during the 2007–2009 financial crisis

Binlei Gong[1] · Robin C. Sickles[2]

Published online: 10 January 2020

Abstract

The paper compares unrestricted and restricted reduced-form estimates of productivity and efficiency performance constructed from non-structural stochastic frontier analysis (SFA) and the structural models of Olley–Pakes (OP), Ackerberg–Caves–Frazer (ACF), Pakes–McGuire (PM), and Midrigan–Xu (MX). These methods are used to estimate changes in firm-level manufacturing productivity in the British Isles before and after the 2007–2009 financial crisis using the Financial Analysis Made Easy (FAME) data set over the period 2005–2012. The empirical results indicate that overall technical efficiency was not impacted to any substantial degree by the financial crisis, according to all models. The empirical results also indicate substantial agreement in the predictions of productivity growth for the three models. The SFA framework (and related DEA approaches) is used internationally to set tariffs in regulated industries. To have SFA, and the OP/ACF/PM/MX models that are more highly leveraged on economic optimizing assumptions, provide comparable estimates of productivity and efficiency change is reassuring. However, it also would suggest that structural approaches may not provide regulators much new information about efficiency and productivity than would less structural approaches such as SFA, while being less transparent and more difficult to justify to regulators as well as to the courts to which regulated firms often turn for relief from tariffs they perceive to be unfair or onerous.

Keywords Productivity · Non-structural and structural models · Stochastic frontier analysis · Olley–Pakes · Pakes–McGuire

JEL D24 · D22 · C23 · G31 · L60 · G01

1 Introduction

The performance of a firm is usually measured by productivity, the ratio of a weighted average of outputs to a weighted average of inputs. Traditional productivity measurement includes labor productivity, capital productivity, and so on. Total Factor Productivity (TFP) accounts for the portion of output not explained by traditionally measured inputs utilized in production (Comin 2010). TFP effectively evaluates overall productivity and has been widely used in the literature (Boucinha et al. 2013; Cummins and Xie 2013; Ball et al. 2014; Curi and Lozano-Vivas 2015; Asmild et al. 2016; Gong 2018a) and in the analysis of growth by statistical agencies across the world.

Non-structural approaches (Perelman 1995; Van Dijk and Szirmai 2011; Gong 2018b) have been extensively employed to estimate productivity. Stochastic Frontier Analysis (SFA) is a well-known method that estimates the average frontier function as well as the level of technical efficiency and its change over time. Neutral technical changes using a SFA approach can also be identified with the panel stochastic frontier model employed in our analysis. Such an approach uses a normalization in which the firm with the highest level of technical efficiency can be identified and the technical efficiency of all other firms can be measured relative to the most efficient firm in the sample at any point in time (Cornwell et al. 1990). Although this non-structural model can estimate the average technical efficiency level, and therefore predict the overall relative TFP loss for each unit

✉ Binlei Gong
gongbinlei@zju.edu.cn

✉ Robin C. Sickles
rsickles@rice.edu

[1] School of Public Affairs, China Academy for Rural Development (CARD), and Academy of Social Governance, Zhejiang University, Zhejiang, China

[2] Department of Economics, Rice University, Houston, TX, USA

due to technical inefficiency, it cannot explain the sources of the inefficiency or TFP loss without a more detailed explanation of the inefficiency residual term. This has been done in a variety of ways, typically by specifying a reduced form model for the exogenous determinants of efficiency.[1]

Productivity defined in this way by a Solow residual is a non-structural concept and cannot be given a structural interpretation without a more formal model in which an explicit set of orthogonality conditions are specified that allow a counterfactual analysis to be undertaken. Our study complements non-structural SFA with two structural models from the literature and compares their restricted reduced-form predictions, using panel data from firms in the British Isles.

The first group of structural models that we consider include Pakes–McGuire (PM) model (Pakes et al. 1993; Pakes and McGuire 1994) and Midrigan–Xu (MX) model (Midrigan and Xu 2014). On the one hand, PM computes the Markov Perfect Nash (MPN) Equilibria of an industry, where a firm's profit is a function of its own level of efficiency as well as the efficiency level of all its competitors. This approach can estimate an average efficiency level by solving both the MPN equilibria and the social planner's problem. On the other hand, MX decomposes the residual and estimates the partial TFP loss due to specific factors (financial friction or financial misallocation, in this case). This approach sets up an economic structure and calculates the efficient allocation (with highest attainable productivity) to satisfy the labor market clearing condition, the asset clearing condition, producer and worker optimization, and the no-arbitrage condition. Both PM and MX require only summary statistics data to solve equilibrium and estimate productivity.

The second group of structural models that we consider to estimate productivity contains the Olley–Pakes method (OP) (Olley and Pakes 1996) and Ackerberg–Caves–Frazer method (ACF) (Ackerberg et al. 2015) of the production function estimation. Analogous to SFA, OP also estimates productivity based on a production function, but the behavioral framework of OP has similarities with the structural PM. However, the estimation algorithm of OP is based on five extra assumptions (see Ackerberg et al. 2015, page 2416), which relies heavily on the structural model in Pakes (1996). Therefore, this paper treats OP as a model somewhere between the non-structural SFA and the structural PM. ACF points out that OP fails to identify the labor elasticity in the first step and should be estimated in the second step. Different from PM and MX, OP and ACF require firm-level data to estimate productivity, which is similar with reduced form methods such as SFA.

Some studies also compare non-structural models and structural models in productivity analysis. Eberhardt and Helmers (2010) focus on the transmission bias problem of the production function and compare dynamic panel estimators (Arellano and Bond 1991; Blundell and Bond 1998) with structural OP and ACF models. Del Gatto et al. (2011) survey many productivity and efficiency methods, including the non-structural stochastic frontier model and the structural OP and ACF models. Their contribution is to choose the right model among available methodologies, especially the difference in macro studies and micro studies.

In our paper, we emphasize that the non-structural stochastic frontier model is the better option as it does not rely on strong identifying assumptions and therefore is more robust. The only way to reasonably make this argument is to base it on empirical evidence, which is not included in Del Gatto et al. (2011). Therefore, the purpose and motivation of these two papers are different. Moreover, our paper also introduces two structural models (Pakes–McGuire model and Midrigan–Xu model) that are not included in Eberhardt and Helmers (2010) and Del Gatto et al. (2011).

We analyze the effect of financial frictions on productivity before and after the 2007–2009 financial crisis using firm-level data for the manufacturing sectors of the British Isles. The overall TFP loss estimate using SFA, OP, ACF and PM will be compared. Our empirical findings indicate that the overall TFP loss was not impacted to any substantial degree by the financial crisis, according to all models. The fact that we find similar results using very different methodologies adds to the confidence that the overall TFP loss ratio was not significantly affected by the financial crisis. Moreover, MX is adopted to estimate productivity loss due to financial friction from the recession, which further supports our findings derived from SFA, OP, ACF and PM.

The remainder of the article is structured as follows. Section 2 introduces the Stochastic Frontier Analysis. Section 3 describes the Pakes–McGuire (PM) and Midrigan–Xu (MX) models. Section 4 presents the Olley–Pakes (OP) and Ackerberg–Caves–Frazer (ACF) models. Section 5 summarizes the assumptions of different non-structural and structural models. Section 6 provides data descriptions. Empirical results derived from the three approaches are presented and compared in Section 7. Section 8 concludes.

2 Stochastic frontier analysis

2.1 Stochastic frontier analysis of cross-sectional data

Aigner et al. (1977) and Meeusen and Van den Broeck (1977) proposed the stochastic frontier production function

[1] The usefulness and importance of SFA (and DEA) methods have passed the market test as they are required to be used in a wide variety of regulatory decision-making settings in Europe and elsewhere (e.g., Bogetoft 2013, Agrell and Bogetoft 2017, and Agrell and Bogetoft 2018).

model of the form:

$$y_i = x_i'\beta + \nu_i - u_i, \qquad i = 1, \ldots, N. \tag{1}$$

In this paradigm the deterministic frontier production function is augmented by a symmetric random error variable, v_i, to account for measurement errors and other sources of non-systematic statistical noise. y_i is the output of firm i, often in logarithms x_i is the vector of inputs often in logarithms (Cobb–Douglas or translog specifications), and u_i is a non-negative random variable representing technical inefficiency (the distance to the frontier).

In most cases v_i is assumed to follow a normal distribution that is independent of each u_i. Both v_i and u_i are often assumed to be uncorrelated with the independent variables x_i. A variety of distributional assumptions are also imposed on u_i. Aigner et al. (1977) assumed u_i to be i.i.d. half-normal random variables and derived the Maximum Likelihood (ML) estimates. Stevenson (1980) introduced a normal truncated specification, while Greene (1990) considered the gamma specification.

2.2 Stochastic frontier analysis of panel data

The stochastic frontier literature in the early 1980s mainly consists of analyses of cross-sectional data. Pitt and Lee (1981) introduced a parametric MLE random effects generalization that allowed a constant firm specific efficiency level to be consistently estimated while Schmidt and Sickles (1984) addressed the difficulties with the canonical stochastic frontier model, including inconsistent firm-specific technical inefficiency estimations, strong assumptions about the distribution of technical inefficiency and statistical noise, and potentially incorrect assumptions that inefficiency is independent of the regressors. Cornwell et al. (1990) provided a variety of estimators that not only solved these problems but also allowed for the firm-specific efficiency levels to vary over time. Their panel stochastic frontier model is

$$\begin{aligned} y_{it} &= \alpha + x_{it}'\beta + \nu_{it} - u_{it} = \alpha_{it} + x_{it}'\beta + \nu_{it}, \\ &i = 1, \ldots, N,\ t = 1, \ldots, T \end{aligned} \tag{2}$$

Cornwell et al. (1990) introduced the within estimator (CSSW), the generalized least squares estimator (CSSG), and a Hausman and Taylor (1981)-type estimator they labeled the Efficient IV estimator that allow for time-variant individual efficiency effects by replacing the firm effect with heterogeneous environmental variables whose effect on efficiency was firm specific. In their empirical example, they assumed that only the coefficients on the time and time-squared variables have such heterogeneous patterns, resulting in a parameterization of the firm effects of α_{it} that took the form of $\alpha_{it} = \theta_{i1} + \theta_{i2}t + \theta_{i3}t^2$. Sickles (2005) later examined various specifications of the time-variant firm effect α_{it} modeled in other research, including $\alpha_{it} = \eta_{it}\alpha_i = \exp[\eta(1-T)]\alpha_i$ (Battese and Coelli 1992), and the general factor Kneip-Sickles-Song (KSS) model $\alpha_{it} = c_{i1}g_{1t} + c_{i2}g_{2t} + \cdots + c_{iL}g_{Lt}$ (Kneip 1994; Kneip et al. 2003),[2] which are widely used SFA approaches in literature (Kumbhakar and Wang 2005; Gong 2019). Instead of choosing one of the SFA models, Gong (2018c) used a jackknife model averaging method to consider CSSG and KSS models simultaneously.

2.3 Productivity and efficiency estimates under different SFA models

We use the Fixed Effects (FIX), Random Effect (RND), Kneip-Sickles-Song (KSS), and Battese–Coelli (BC) non-structural specifications to estimate efficiency and productivity. The first two are time-invariant estimators, and the last two are time-varying effects estimators. We assume a Cobb–Douglas stochastic frontier function with constant returns to scale as the main model, because many studies (Burnside 1996; Basu and Fernald 1997; Gong 2020) of manufacturing have found it difficult to reject such a restriction. However, in robustness checks, we also assume a Cobb–Douglas stochastic frontier function without constant returns to scale and a translog stochastic frontier function, respectively. Thus the main model estimates:

$$\begin{aligned} &\log(Y_{it}) = \alpha_t + \beta_1\log(K_{it}) + \beta_2\log(L_{it}) + \nu_{it} - u_{it} \quad \text{s.t. } \beta_1 + \beta_2 = 1 \\ &\quad \Rightarrow \log(Y_{it}) = \alpha_t + \beta_1\log(K_{it}) + (1-\beta_1)\log(L_{it}) + \nu_{it} - u_{it} \\ &\quad \Rightarrow \log(Y_{it}) - \log(L_{it}) = \alpha_t + \beta_1((\log(K_{it}) - \log(L_{it})) + \nu_{it} - u_{it} \end{aligned} \tag{3}$$

In a single output world with constant returns to scale, the TFP for firm i at time t is equal to $\exp(\alpha_t - u_{it} + v_{it})$, which decomposes to the frontier ($\exp(\alpha_t + v_{it})$) and firm-specific efficiency ($\exp(-u_{it})$). Therefore, the growth in TFP is equal to the sum of the frontier shifts over time that measures the technical changes and the distance of a firm to the frontier that measures efficiency changes. In the FIX and RND models, $\alpha_t = \alpha + \delta t$ and $u_{it} = u_i$, which implies the degree of technical progress is fixed at δ and the firm-specific efficiency doesn't change over time. In the BC model, $\alpha_t = \alpha + \delta t$ and $u_{it} = \exp(-\eta(t-T)) \cdot u_i$, where $u_{it} \sim N^+(\mu, \sigma_\mu^2)$ is a truncated normal distribution. Therefore, the degree of technical progress is still fixed at δ, but the firm-level efficiency is time-variant. However, the efficiency changes at the same speed across firms and time in the BC model, as the growth rate $\exp(-\eta)$ is fixed. In the KSS model, $u_{it} = \sum_{r=1}^{L}\theta_{ir}g_r(t)$ where $g_1(t), \ldots, g_L(t)$ are the

[2] Factor models related to Kneip-Sickles-Song model has been studied in Forni et al. (2000), Bai and Ng (2002), Stock and Watson (2002), Pesaran (2006), and Bai (2009).

basis functions and $\theta_{i1}, \ldots, \theta_{iL}$ are the coresponding parameters. The change in firm-level efficiency is semiparamatrically estimated, which is both firm-variant and time-variant. The term α_t quantifies a general mean process to ensure identifiability and represents the degree of technical progress in the KSS model.

The equation $\exp(\alpha_{it}) = \exp(\alpha_t - u_{it} + v_{it})$ provides the individual TFP level for the i-th firm at time t, which applies to all SFA models. Assuming that the largest individual TFP is the efficient level of TFP, the average level of TFP loss in percentage can be derived, which indicates the average level of efficiency in the economy.

$$\begin{aligned}
&\text{TFP loss ratio}\\
&= 1 - \frac{TFP}{TFP^e} \approx 1 - \left(\frac{1}{T}\sum_{t=1}^{T}\left(\frac{\frac{1}{N}\sum_{i=1}^{N}\exp(\alpha_t - u_{it} + \nu_{it})}{\exp(\alpha_t)}\right)\right)\\
&\approx 1 - \left(\frac{1}{T}\frac{1}{N}\sum_{t=1}^{T}\sum_{i=1}^{N}\exp(-u_{it})\right) = 1 - \text{Average TE}
\end{aligned} \tag{4}$$

where TFP^e is the efficient level of total factor productivity and TFP is the average level of total factor productivity. The ratio $\frac{TFP}{TFP^e}$ equals the average technical efficiency (TE) that is typically used in the stochastic analysis. This study can thus estimate the average technical efficiency using the stochastic frontier analysis and then derive the TFP loss ratio. The degree of appropriable technical progress accessible by all firms is not taken into consideration when calculating our TFP loss measure as α_t is common to all firms and as such it appears in both the numerator and denominator of Eq. (4).

3 Pakes–McGuire model and Midrigan–Xu model

The first structural model that we consider is the Pakes–McGuire (PM) model (Pakes et al. 1993; Pakes and McGuire 1994). This model computes the MPN equilibria (Maskin and Tirole 1988a, 1988b) generated under the constraints of Ericson and Pakes (1992), where a firm's profit is a function of its own level of efficiency as well as the efficiency level of all its competitors. This approach can estimate an average efficiency level by solving both the MPN equilibria and the social planner's problem. Detailed information of PM can be found in Appendix 1.

The second structural model is the model of Midrigan and Xu (MX) (Midrigan and Xu 2014), which is a more up-to-date structural model that estimates productivity. MX decomposes the residual and estimates the partial TFP loss due to specific factors (financial friction or financial misallocation, in this case). This approach sets up an economic structure and calculates the efficient allocation (with highest attainable productivity) to satisfy the labor market clearing condition, the asset clearing condition, producer and worker optimization, and the no-arbitrage condition. The factors of interest are introduced into the model along with borrowing constraints, which can lead to a lower possible TFP, and therefore cause TFP loss. Detailed information of MXM can be found in Appendix 2.

4 Olley–Pakes and Ackerberg–Caves–Frazer models

Compared with frontier analysis, conventional production function analysis can also be used to estimate efficiency and productivity growth using the Olley and Pakes (1996) approach. However, it is an average production function rather than a frontier production function that is estimated in this setting. Since the production function estimates may suffer from simultaneity, as the efficiency/productivity levels are known to firms when they decide their input utilizations but are unobservable to economists, instruments for endogenous inputs are also needed. For example, profit maximizing firms may cut headcounts when they see negative productivity shocks. This endogeneity problem can lead to biased OLS estimates of the production function and inaccurate TFP levels. Of course the use of instrumental variables or control function modifications of the canonical SFA model introduced over 40 years ago (Aigner et al. 1977) is an available option for the SFA model as well (see, e.g., Amsler et al. 2016 and Kutlu 2018). Moreover, SFA treatments that utilize panel fixed effects estimators, such as those in the Schmidt and Sickles (1984) and the Cornwell et al. (1990) model used in generating our SFA results, address the potential correlation of efficiency effects and input decisions.

Olley and Pakes (1996) established an approach to deal with the simultaneity bias by using investment to proxy for the unobserved productivity/efficiency shock. In this model, the behavioral framework, including the entry and exit decisions, the investment and capital accumulation equations, and the assumption of Markov perfect Nash equilibrium, are very similar to those in the PM. Firms in the OP model make labor and investment decisions to maximize the net present value of future profits. Suppose the production function follows a Cobb–Douglas formation

$$y_{it} = \beta_k k_{it} + \beta_l l_{it} + w_{it} + \epsilon_{it}, \tag{5}$$

where y_{it} is the log of output, k_{it} is the log of capital, and l_{it} is the log of labor. w_{it} measures productivity/efficiency shocks and ε_{it} is the usual idiosyncratic disturbance term. As summarized in Ackerberg et al. (2015), there are five assumptions in the OP model. These involve the firm's information set of productivity shocks, first order Markov evolution of productivity shocks, timing of input decisions, the form of the investment function, and the strict monotonicity of productivity and levels of investment. The fourth assumption states that investment is a function of capital and productivity shocks: $i_{it} = f_t(k_{it}, w_{it})$. The last assumption of strict monotonicity of productivity and investment levels, derived from Pakes (1996), results in the function $f_t(k_{it}, w_{it})$ that is strictly increasing in w_{it}, indicating that the inverse function $w_{it} = f_t^{-1}(i_{it}, k_{it})$ exists. Substituting this inverse function into the production function in Eq. (5) yields $y_{it} = \beta_k k_{it} + \beta_l l_{it} + f_t^{-1}(i_{it}, k_{it}) + \epsilon_{it} = \beta_l l_{it} + \phi_t^{-1}(i_{it}, k_{it}) + \epsilon_{it}$. where the parameter β_l is point identified. OP treat ϕ_t^{-1} as a non-parametric function in order obtain $\hat{\beta}_l$ when they estimate this last function. However, since $\beta_k k_{it}$ and $f_t^{-1}(i_{it}, k_{it})$ are both included in the non-parametric function ϕ_t^{-1}, β_k cannot be estimated without a further step. In the second step, OP uses the timing of capital decision (the third assumption) to point identify and estimate β_k. The innovation ξ_{it} is defined as the difference between the expected productivity/efficiency of the next period and realized productivity. The conditional expectation of innovation at time t is zero given information at time $t-1$, i.e. $E[\xi_{it}|I_{it-1}] = 0$. The information set I_{it-1} includes k_{it}, since current capital equals previous capital less depreciation plus previous investment and thus $E[\xi_{it}|k_{it}] = E[\xi_{it}k_{it}] = 0$. This orthogonality condition identifies β_k.

Next, given an estimate of β_k, OP first derive the level of efficiency/productivity as $w_{it}(\beta_k) = \widehat{\phi}_t(i_{it}, k_{it}) - \beta_k k_{it}$, where $\widehat{\phi}_t(i_{it}, k_{it})$ has been estimated in the first step. Then the innovation is $\xi_{it}(\beta_k) = w_{it}(\beta_k) - \hat{\chi}(w_{it-1}(\beta_k))$ where $\hat{\chi}(w_{it-1}(\beta_k))$ accounts for the fitted value of the conditional expectation from the non-parametric regression. Finally, OP use Generalized Method of Moments (GMM) to solve $\min_{\beta_k}\left[\frac{1}{T}\frac{1}{N}\sum_t \sum_t \xi_{it}(\beta_k)k_{it}\right]$, which satisfies the orthogonality $E[\xi_{it}|k_{it}] = E[\xi_{it}k_{it}] = 0$.

To summarize, OP derive β_l in the first step and then estimate β_k as well as w_{it} in the second step. The productivity/efficiency term w_{it} in OP is comparable with the productivity/efficiency term α_{it} in SFA. Finally, OP method also uses Eq. (4) to calculate the TFP loss ratio.

Compared with OP, the Levinsohn-Petrin (LP) model uses an intermediate input rather than investment to invert and solve for w_{it}. This alternative inversion is because investment is often lumpy and is often zero when data are disaggregated by firm/establishment. This of course also puts into question the strict monotonicity assumption in Olley–Pakes. Thus in LP we have $m_{it} = f_t(k_{it}, w_{it})$ and $w_{it} = f_t^{-1}(m_{it}, k_{it})$, where m_{it} is the log of the intermediate input. Substituting this into the production function, we have $y_{it} = \beta_k k_{it} + \beta_l l_{it} + \beta_m m_{it} + f_t^{-1}(m_{it}, k_{it}) + \epsilon_{it}$. Similar to OP, we can identify β_l in the first step regression (f_t^{-1} is again a non-parametric function). In this model, $\widehat{\phi}_t(m_{it}, k_{it}) = \beta_k k_{it} + \beta_m m_{it} + f_t^{-1}(m_{it}, k_{it})$. In the second step, we have two parameters to estimate. The first parameter, β_k, can be analogously identified as in the OP, where innovation is orthogonal to capital. For β_m, LP adds the moment that innovation is orthogonal to lag intermediate input. Econometrically, those two moments are: $E[\xi_{it}(\beta_k, \beta_m)|k_{it}] = 0$ and $E[\xi_{it}(\beta_k, \beta_m)|m_{it-1}] = 0$. The reduced-form regression (GMM) of the second step is similar to that in OP.

The Ackerberg–Caves–Frazer (ACF) model argues that the first stage of the OP and LP models correctly identify β_l only under several specific conditions. Otherwise, OP and LP fail to provide consistent estimates of β_l. The main difference of ACF compared with OP and LP is that ACF invert the conditional instead of the unconditional input demand function to control for unobserved productivity w_{it}. Different from OP and LP that estimate β_l in the first stage, all the coefficients are estimated in the second stage in ACF model. The detailed estimation strategy can be found in Ackerberg et al. (2015). All these three models, including OP, LP, and ACF, are control function approaches (Mollisi and Rovigatti 2017).

5 Assumptions of non-structural and structural models

This section summarizes and compares the assumptions of non-structural SFA models and structural OP and ACF models.[3] To conclude, non-structural SFA model requires many less assumptions than the structural OP and ACF models.

Stochastic frontier methods include but are not limited to the following assumptions (A.1–A.3): A.1 There is a production frontier for all firms in the same industry, which represents the highest attainable producitiy under current technology; A.2 The gap between the frontier and actual production of a companies is explained by technical inefficiency; A.3 Some SFA estimators (e.g., BC estimator introduced in this article) assume a parametric form for the efficiency and the disturbance; however, other SFA estimators (e.g., FIX and RND estimators introduced in this article and CSS estimator (Cornwell et al. 1990)) are semiparametric efficient under the assumptions of correlated random effects (Mundlak 1978; Chamberlain 1984) as shown in Schmidt and Sickles (1984) and by extension in Cornwell et al. (1990); A.4: Endogenous inputs are addressed in the panel estimators we utilize by the Schmidt and Sickles and Corwell et al. estimators to address the presence of firm effects that are potentially correlated with

[3] For the more complicated PM and MX models, please read Appendix 1 and Appendix 2 for details.

the efficiency terms. These can be static or time-varying effects. Endogenous input choice that is correlated with the usual disturbance term is addressed via standard instruments based on lagged input values as well as input prices, time trends, and other available exogenous variables in the model. The validity of the instruments are tested via the methods introduced by Amsler et al. (2016).

The Olley–Pakes model is assumed to include but is not limited to the following assumptions (B.1–B.6): B.1 Company i at time t knows its current and past productivity shocks but does not know future productivity shocks (i.e., $I_{it} = (w_{i1}, w_{i2}, \ldots, w_{it})$. The transitory shocks ε_{it} follow $\mathrm{E}[\varepsilon_{it}|I_{it}] = 0$; B.2 Productivity shocks at time $t+1$ based on the distribution $p(w_{it+1}|I_{it}) = p(w_{it+1}|w_{it})$, are known to firm i and stochastically increasing in w_{it}; B.3 The state variables at time t, including capital, evolve based on the investment policy function that is determined in the previous period (time $t-1$), whereas free variables, including labor inputs, are selected in the current period (time t) after the firm observes its productivity shocks; B.4 Company i at time t makes investment decisions according to $i_{it} = f_t(k_{it}, w_{it})$; B.5 the function $f_t(k_{it}, w_{it})$ is strictly increasing in w_{it}; and B.6 data generating process (DGP) of the production process is one of the three DGPs mentioned in Ackerberg et al. (2015).

The Ackerberg–Caves–Frazer model is assumed to include but is not limited to the following assumptions (C.1–C.5): C.1 Company i at time t knows its current and past productivity shocks but does not know future productivity shocks (i.e., $I_{it} = (w_{i1}, w_{i2}, \ldots, w_{it})$. The transitory shocks ε_{it} follow $E[\varepsilon_{it}|I_{it}] = 0$; C.2 Productivity shocks at time $t+1$ based on the distribution $p(w_{it}+1|I_{it}) = p(w_{it+1}|w_{it})$, which is known to firm i and stochastically increasing in w_{it}; C.3 The state variables at time t, including capital, evolve based on the investment policy function that is determined in the previous period (time $t-1$), whereas free variables, including labor inputs, are chosen at time $t-1$, time t, or time $t-b$ (which $0<b<1$); C.4 Company i at time t make investment decision according to $i_{it} = f_t(k_{it}, l_{it}, w_{it})$;[4] and C.5 the function $f_t(k_{it}, l_{it}, w_{it})$ is strictly increasing in w_{it}.

6 Data

As one of the developed countries with high financial risk (Derbali 2016), much attention has been paid to the economic performance of U.K. before and after the 2007–2009 financial crisis. Our data come from a commercial software package called Financial Analysis Made Easy (FAME)[5] and cover the years 2005–2012. FAME contains comprehensive information, including balance sheets and profit and loss accounts, for the firms in the UK and Ireland. There are some advantages of FAME data. Each registered firm in the UK is required to provide accounting and other data about their operations to an executive agency of the Department of Trade and Industry know as Companies House, which are then made available in FAME (Graham et al. 2010). Since it covers the entire population of the registered firms including the non-stock market listed firms, Draca et al. (2011) pointed out that FAME can more comprehensively reflect the overall situation of all sized firms than other datasets such as National Minimum Wage (NMW). Moreover, accounting regulations in the UK require private firms to report significantly more accounting information than some other countries. For example, even publicly traded firms may not provide remuneration information in the US, which is available for most firms in FAME.

Because of the rich information and aforementioned advantages, many scholars have used the financial data from FAME in different fields of economics. Girma et al. (2008) investigated the two-way relationship between R&D and export activity using the sales, employment, and wage information from FAME. Guariglia and Mateut (2010) studied the links between firm's global engagement status and their financial health, where the profit and loss and balance sheet data in FAME are utilized. Draca et al. (2011) evaluated the impact of minimum wage on firm profitability in the UK using FAME's profit and remuneration information.

More importantly, the FAME dataset has been used extensively in empirical work involving productivity analysis (Eberhardt and Helmers 2010). In order to study the relationship between the density of economic activity and productivity in the UK, Graham (2007) used FAME data and a production function to estimation TFP. Harris and Li (2008) considered the contribution of exporters to aggregate productivity growth using FAME database. Faggio et al. (2010) also collected data from FAME to estimate TFP for UK firms, which enabled them to further study the evolution of inequality in productivity.

Since FAME covers the entire population of registered UK firms, some other databases use data from FAME to augment the coverage of their data. For example, the Oxford Firm Level IP (OFLIP) merges FAME and IP activity of firms in the form of patents and trade-marks, which is used in Helmers and Rogers (2010), Helmers et al. (2011), and Helmers and Rogers (2011). AMADEUS, a pan-European firm-level database, contains a subset of firms

[4] This is the assumption for ACF correction to OP model. We can replace investment with intermediate input for ACF correction to Levinsohn–Petrin (LP) model when intermediate input rather than investment is used as proxy variable.

[5] http://www.bvdinfo.com/en-gb/our-products/company-information/national-products/fame

contained in FAME, which is also utilized by some scholars (e.g., in Benfratello and Sembenelli 2002).

Each firm's balance sheet provides total asset, liabilities, and shareholder funds information, while the profit and loss account provides operating profits, depreciation, amortization, impairment, remuneration, directors' remuneration, and the number of employees. This information allows us to construct a more accurate measure of the output (Y) and inputs (L and K).

The labor expenditure is the sum of remuneration and directors' remuneration. Remuneration includes wages and salaries, social security cost, pension costs, and other staff costs, while directors' remuneration includes director's fees, pension contributions, and other compensations. Labor quantity is defined as the number of employees.

Capital expenditure is defined as the sum of depreciation, amortization, and impairment. Capital services are based on "capital employed". Employed capital is the total assets less current liabilities. These are the values of the assets that contribute to a firm's ability to generate revenues (liquidity), including both sunk cost (not used in production) and capital (used in production). We assume the same ratio between sunk cost and capital across firms. We use an 8 percent depreciation rate to adjust capital expenditures, which is recommended for the UK during our sample period by Chadha et al. (2016). Total sunk cost for all of the firms is total capital employed minus the total capital. This provides us with the average ratio between sunk cost and capital, which can be applied to estimate the capital in each firm.

In many productivity analyses (McGuckin et al. 1992; Timmer and Voskoboynikov 2014), the output is measured by value added. One way to calculate value added is to use revenue minus cost of intermediate goods if we can identify the net expenditures on intermediate inputs, which is denoted as the subtraction method. Otherwise, we can use the addition method, which is to sum profit, depreciation cost and labor cost. Since FAME does not have complete intermediate input information, we use the second approach. We calculate value added as the sum of operating profit, labor expenditure, and capital expenditure, which is introduced in Griffith et al. (2006) who use this summation method to generate value added and treat it as the output in a conventional Cobb–Douglas production function for manufacturing firms in the UK.

All series are real prices in 2005 British Pounds (GBP). We deflate value-added, labor cost, and capital cost using the Producer Price Index (PPI) and Consumer Price Index (CPI): PPI and CPI of Ireland for firms in the Republic of Ireland and the PPI and CPI of the UK for firms in England, Scotland, Wales, North Ireland, and British Crown dependencies.

Using the FAME data, Harris and Li (2008) found huge differences between manufacturing and non-manufacturing sectors, and hence calculated the productivity for the two groups separately. In this paper, we utilize only firms in the manufacturing sectors of the British Isles, rather than firms in all sectors, in order to minimize heterogeneity in production processes. Missing observations, obvious reporting errors, and outliers (largest and smallest 0.5%) are excluded from our sample, leaving us with $4{,}814 \times 8 = 38{,}512$ firm-year observations over an eight-year period from 2005 to 2012.

Table 1 shows the sample statistics of the FAME data we utilize in our empirical analysis. Average output is £23 million before the crisis and £26 million after the crisis. Capital on average increases from £31 million before the crisis to £35 million after the crisis, while its expenditure increases from £2.49 million to £2.82 million. The average number of employees also increases from 447 to 454 between those two periods, while labor costs increase from £14.04 million to £14.82 million. These statistics point to an increase in output, capital, and labor of 11, 13, and 2%, respectively, indicating that firms in the British Isles were using more capital input in their portfolios after the financial crisis.

7 Quantitative analysis

7.1 SFA/OP/ACF model estimates of overall TFP loss

Table 2 compares the estimation results for the stochastic frontier models and control function approaches for the

Table 1 Summary statistics

Variable	Pre-financial crisis (2005–2007)				Post-financial crisis (2008–2012)			
	Mean	S.D.	Min	Max	Mean	S.D.	Min	Max
Output (Y) (£ million)	23.44	261.0	0.081	13,900	26.04	294.2	0.03	14,700
Capital (K) (£ million)	31.15	349.0	0.007	14,100	35.32	442.8	0.014	17,300
Labor (L)	447	3153	2	103,401	454	3331	2	107,046
Capital expenditure (δK)	2.49	22.7	0	1123	2.82	31.6	0	2047
Labor expenditure (WL)	14.04	114.4	0.01	5484.2	14.82	129.9	0.01	6478.7

Notes: The sample size of pre-financial crisis is 14,442 and the sample size of post-financial crisis is 24,070 pre-crisis period (2005–2007). Table 3 provides similar comparisons but for the post-crisis period (2008–2012). It is worth noting that we use the method in Amsler et al. (2016) to check the endogeneity of the inputs where input prices and lagged input quantities are adopted as instruments. The result shows that both labor and capital in our dataset are exogenous. The first four columns of Tables 2 and 3 report the estimations derived from four SFA models (FIX, RND, KSS and BC) discussed in Section 2. The fifth columns of Tables 2 and 3 report the OLS estimates of the conventional production function as the benchmark, whereas the sixth and seventh column reports the estimation of OP and ACF models introduced in Section 4, respectively. The first two rows in Tables 2 and 3 give the estimated coefficient and standard error of the simple time trend, respectively. All estimators show that the technology change increases TFP in the first period, but decreases TFP after the financial crisis. The second two rows are the coefficient estimation and standard error of capital input, respectively.

Tables 2 and 3 also reports average technical efficiency, as well as the TFP loss by different models for the pre- and post-crisis periods, which can be directly derived after we estimate the average technical efficiency using Eq. (4). The estimated TFP losses based on estimates from the four SFA models and the OP model are fairly robust, and the average is 50% (equal to an average 50% technical efficiency) before the financial crisis. These TFP losses by different models are all stable after the recession and remain at 50% (equal to 50% average efficiency).

This paper adopts four robustness checks. Firstly, Table 4 compares the distribution of technical efficiency under SFA/OP/ACF models. The results show that different models not only provide robust average technical efficiency, but also robust 10, 25, 50, 75, and 90% quantile technical efficiency estimates. Secondly, Table 5 reports the TFP loss ratio year by year, which is fairly robust and further confirms the stable TFP losses after the financial crisis. Thirdly, we check the robustness of the results under different functional form assumptions. Table 6 compares the TFP loss ratio the pre-crisis period (2005–2007) under the main model where the production function is Cobb–Douglas with constant returns to scale, as well as the ones estimated under Cobb–Douglas and translog formulations without constant returns to scale. Table 7 provides similar estimations for the post-crisis period (2008–2012). The results from the different specifications are quite robust. Finally, although SFA/OP/ACF models generate similar average technical efficiency and productivity loss in FAME data, whether similar results can be yielded in other dataset is unknown. To this end we examine a sectorial subsample of Chilean firm-level

Table 2 Stochastic frontier and control function estimators for pre-financial crisis (2005–2007)

	Stochastic frontier approach				Control function approach		
	FIX	RND	KSS	BC	OLS	OP	ACF
Time trend	0.020***	0.016***	–	−0.014	0.013***	0.011**	0.015***
	(0.002)	(0.002)		(0.016)	(0.005)	(0.005)	(0.006)
Capital	0.174***	0.244***	0.206***	0.244***	0.300***	0.359***	0.305***
	(0.008)	(0.006)	(0.009)	(0.006)	(0.004)	(0.039)	(0.040)
Labor	0.826***	0.756***	0.794***	0.756***	0.700***	0.641***	0.695***
	(0.008)	(0.006)	(0.009)	(0.006)	(0.004)	(0.044)	(0.016)
Average TE	0.485	0.507	0.485	0.524	0.535	0.514	0.529
TFP loss ratio	0.515	0.493	0.515	0.476	0.465	0.486	0.471

Significant at: *10, **5 and ***1 percent; Standard error in parentheses

Table 3 Stochastic frontier and control function estimators for post-financial crisis (2008–2012)

	Stochastic frontier approach				Control function approach		
	FIX	RND	KSS	BC	OLS	OP	ACF
Time trend	−0.007***	−0.007***	–	−0.160***	−0.008***	−0.007***	−0.008***
	(0.001)	(0.001)		(0.007)	(0.002)	(0.003)	(0.002)
Capital	0.224***	0.273***	0.286***	0.252***	0.339***	0.394***	0.405***
	(0.006)	(0.005)	(0.008)	(0.005)	(0.003)	(0.031)	(0.047)
Labor	0.776***	0.727***	0.714***	0.748***	0.661***	0.606***	0.595***
	(0.006)	(0.005)	(0.008)	(0.005)	(0.003)	(0.028)	(0.046)
Average TE	0.497	0.504	0.490	0.519	0.512	0.492	0.495
TFP loss ratio	0.503	0.496	0.510	0.481	0.488	0.508	0.505

Significant at: *10, **5 and ***1 percent; Standard error in parentheses production data[6] in Table 8 to check if SFA/OP/ACF models still provide robust results. The estimated average technical efficiency and productivity losses are again robust across approaches.

Table 9 reports the average TFP loss for each of the 24 sectors in UK manufacturing according to UK SIC 2007 code. The average TFP loss is the mean value of four SFA estimates (FIX/RND/KSS/BC). During the sample period, 14 manufacturing sectors suffered from less than a 50% TFP loss, whereas 10 other manufacturing sectors suffered more than 50% TFP loss. On the one hand, manufacture of coke and refined petroleum products (SIC code 19), manufacture of basic pharmaceutical products and pharmaceutical preparations (SIC code 21), as well as repair and installation of machinery and equipment (SIC code 33) are the three sectors that experienced the least TFP loss. On the other hand, manufacture of textiles (SIC code 13), manufacture of wood and of products of wood and cork, except furniture; manufacture of articles of straw and plaiting materials (SIC code 16), and manufacture of furniture (SIC code 31) encountered the most TFP loss. To summarize, significant TFP losses are broad-based across sectors, since 23 out of the 24 sectors experienced more than 40% TFP loss. Bughin et al. (2018) and Tenreyro (2018) both point out that the manufacturing sector has an outsize impact on the productivity slowdown in the UK. Riley et al. (2018) believe that the movement of jobs out of productive manufacturing into other less productive sectors in the UK is an important reason for the productivity slowdown at the aggregate level. Our results of broad-based TFP losses in manufacturing is consistent with the findings in Bughin et al. (2018). This productivity slowdown can lead to less employment in manufacturing and consequently can result in the productivity slowdown for the whole economy as described in Riley et al. (2018).

Table 4 Distribution of technical efficiency

	FIX	RND	KSS	BC	OLS	OP	ACF
Pre-financial crisis (2005–2007)							
10% quantile	0.374	0.396	0.393	0.417	0.436	0.410	0.429
25% quantile	0.458	0.482	0.452	0.502	0.484	0.461	0.478
50% quantile	0.485	0.508	0.488	0.527	0.531	0.511	0.526
75% quantile	0.564	0.591	0.567	0.609	0.583	0.565	0.578
90% quantile	0.641	0.675	0.644	0.683	0.638	0.622	0.634
Post-financial crisis (2008–2012)							
10% quantile	0.382	0.393	0.371	0.389	0.409	0.382	0.386
25% quantile	0.469	0.478	0.462	0.449	0.461	0.438	0.442
50% quantile	0.498	0.504	0.487	0.532	0.510	0.490	0.494
75% quantile	0.578	0.588	0.574	0.586	0.561	0.544	0.548
90% quantile	0.651	0.655	0.651	0.634	0.617	0.601	0.604

Table 5 TFP loss ratio by year in various models

	FIX	RND	KSS	BC	OLS	OP	ACF
2005	0.515	0.493	0.517	0.470	0.473	0.489	0.478
2006	0.515	0.493	0.502	0.475	0.463	0.487	0.469
2007	0.515	0.493	0.517	0.473	0.460	0.484	0.467
2008	0.503	0.496	0.510	0.515	0.477	0.498	0.501
2009	0.503	0.496	0.511	0.498	0.497	0.522	0.521
2010	0.503	0.496	0.510	0.480	0.493	0.510	0.508
2011	0.503	0.496	0.500	0.463	0.470	0.493	0.488
2012	0.503	0.496	0.513	0.476	0.502	0.519	0.509

7.2 Pakes–McGuire algorithm based on overall TFP loss

Pakes Gowrisankaran and McGuire (1993) and Pakes and McGuire (1994) defined several constants in their algorithm. The same parameters as the original paper, except for three in the profit function, are utilized. We also allow the profit function to change before and after the financial crisis. PM provides the outputs, including the efficiency level of all the active firms in every period, the average lifespan of firms, the average investment and profit in one period, and so on. To this end, we calibrate the three parameters in the profit function (D, f, and γ) by requiring that the model provides similar statistics with FAME data, including the average lifespan of firms and the investment-to-profit ratio. Then, the average efficiency level can be estimated.

We follow Pakes et al. (1993) and Pakes and McGuire (1994) by assuming an economy that starts with one firm and that the efficiency level of firms can range from 0 to 19. Pakes et al. assume new firms enter with an efficiency level of 4. In our analysis we calculate the average efficiency of the newly established firms in the FAME data, which is 0.45 using SFA and therefore assume new firms in the UK enter with

Table 6 Estimation of TFP loss ratio under various specifications for pre-financial crisis (2005–2007)

	Stochastic frontier approach					Control function approach		
	FIX	RND	KSS	BC	SFAs average	OLS	OP	ACF
Cobb–Douglas with CRS	0.515	0.493	0.515	0.476	**0.500**	0.465	0.486	0.471
Cobb–Douglas without CRS	0.556	0.486	0.539	0.470	**0.513**	0.461	0.480	0.470
Translog without CRS	0.535	0.459	0.565	0.459	**0.505**	0.470	0.489	0.485

Table 7 Estimation of TFP loss ratio under various specifications for post-financial crisis (2008–2012)

	Stochastic frontier approach					Control function approach		
	FIX	RND	KSS	BC	SFAs average	OLS	OP	ACF
Cobb–Douglas with CRS	0.503	0.496	0.510	0.481	**0.498**	0.488	0.508	0.505
Cobb–Douglas without CRS	0.533	0.502	0.509	0.489	**0.508**	0.486	0.499	0.506
Translog without CRS	0.529	0.472	0.575	0.511	**0.522**	0.480	0.508	0.487

Table 8 Stochastic frontier and control function estimators for Chilean firm data

	Stochastic frontier approach			Control function approach		
	FIX	RND	BC	OLS	OP	ACF
Time trend	0.017*** (0.003)	0.018*** (0.003)	0.016* (0.009)	0.019*** (0.004)	0.028*** (0.006)	0.019 (0.015)
Capital	0.226*** (0.027)	0.240*** (0.027)	0.238*** (0.026)	0.274*** (0.042)	0.291*** (0.020)	0.296*** (0.013)
Labor	0.561*** (0.013)	0.525*** (0.012)	0.528*** (0.012)	0.395*** (0.012)	0.411*** (0.053)	0.401*** (0.058)
Water	−0.111*** (0.022)	−0.103*** (0.022)	−0.104*** (0.022)	−0.085** (0.038)	−0.051** (0.024)	−0.076*** (0.013)
Electricity	0.323*** (0.012)	0.338*** (0.011)	0.337*** (0.011)	0.415*** (0.013)	0.349*** (0.043)	0.379*** (0.031)
Average TE	0.776	0.775	0.774	0.769	0.766	0.768
TFP loss ratio	0.224	0.225	0.226	0.231	0.234	0.232

Significant at: *10, **5 and ***1 percent; Standard error in parentheses. Besides labor and capital, the companies in Chilean firm data also use two intermediate inputs including water and electricity. Similar with the major model in Tables 2 and 3, this table assumes Cobb–Douglas production function with constant returns to scale. This table does not calculate KSS estimator, since the Chilean firm data is an unbalanced dataset

efficiency levels of 9 (i.e. 19*0.45). We also looked at the stability of our results to the assumption that the economy begins with one firm by starting with multiple firms and found our results to be quite robust to varying this initial condition.[7]

All of these statistics are computed by simulating such an economy. We also use the value function, investment, and entry/exit decisions to evaluate the optimal policies and update the industry structure. We have separate but similar programs to evaluate the statistics for the MPN equilibrium and for the social planner's problem before and after the crisis, respectively. The industry is simulated 10,000 times and the average efficiency level is the mean of all the active firms in those 10,000 periods.

Based on FAME data, the average lifespan of firms is 17.4 years before and 16.1 years after the crisis. Goodridge et al. (2012) provide the total annual investment in the UK from 2005 to 2011. The annual Gross Value Added (GVA) of the UK is available in the Regional Gross Value Added report from the Office for National Statistics (ONS). These two datasets provide the investment-to-value added ratio. FAME data provides the value added-to-profit ratio. Therefore, the average investment-to-profit ratio can be derived, which calculates to 0.582 before and 0.557 after the recession in 2008.

[7] The SFA approaches in Section 7.1 also give an average efficiency of 44.6% for the newly established firm in 2005 (the entrants) in the British Isles. Therefore, this paper sets a corresponding efficiency level of 9 (out of 19) in the PM approach for new firms.

Table 10 presents the parameters used in the PM, as well as the outputs. The average efficiency level before and after the crisis is 6.85 and 7.03, respectively. Since the highest efficiency level witnessed is 13 in both periods, this study predicts a 47.3 and 45.9% TFP loss ratio before and after the crisis, respectively. As a comparison, the SFA and OP/ACF results indicated around 46–51% TFP loss ratio for both before and after the recession. Therefore, the estimations of the overall TFP loss ratio in all SFA/OP/ACF/PM models are robust around 45–51%, and the change in TFP loss ratio after the recession is not significant.

7.3 Midrigan–Xu model based TFP loss

This model groups the parameters into two categories. The first category includes parameters that determine the process for entrepreneurial productivity, as well as the size of the financing frictions. We calibrate these parameters by requiring that the model accounts for the salient features (Part A in Table 11) of the FAME data. The second category includes preference and technology parameters that are difficult to directly identify using the FAME data. We assign these parameter values as follows: (1) labor elasticity (α) is estimated using the FAME data; (2) span of control (η), discount factor ($\beta(1+\mu)^{-1}$), persistence of workers in the unemployed state (λ_0),[8] and relative efficiency in modern sector ($(1+\eta)\phi$) are similar to the ones in Midrigan and

[8] The probability of unemployed workers staying unemployed.

Table 9 TFP loss ratio for different sectors

SIC code	Sector	TFP Loss
19	Mfg. of coke and refined petroleum products	0.349
21	Mfg. of basic pharmaceutical products and pharmaceutical preparations	0.417
33	Repair and installation of machinery and equipment	0.424
26	Mfg. of computer, electronic and optical products	0.443
28	Mfg. of machinery and equipment n.e.c.	0.447
20	Mfg. of chemicals and chemical products	0.458
11	Mfg. of beverages	0.472
18	Printing and reproduction of recorded media	0.472
32	Other manufacturing	0.485
24	Mfg. of basic metals	0.491
27	Mfg. of electrical equipment	0.491
12	Mfg. of tobacco products	0.495
25	Mfg. of fabricated metal products, except machinery and equipment	0.497
30	Mfg. of other transport equipment	0.497
14	Mfg. of wearing apparel	0.501
29	Mfg. of motor vehicles, trailers and semi-trailers	0.514
23	Mfg. of other non-metallic mineral products	0.518
17	Mfg. of paper and paper products	0.528
22	Mfg. of rubber and plastic products	0.535
15	Mfg. of leather and related products	0.545
10	Mfg. of food	0.547
31	Mfg. of furniture	0.551
16	Mfg. of wood and of products of wood and cork, except furniture; Mfg. of articles of straw and plaiting materials	0.570
13	Mfg. of textiles	0.570

Xu (2014); (3) capital depreciation (δ) follows UK's average level as discussed in Section 6; (4) growth rate (γ) follows the UK's average level discussed in the first category (Part A in Table 11), which is calculated using the FAME data; and (5) persistence of workers in the employed state (λ_1)[9] so that λ_0 and λ_1 guarantee that the fraction of workers who supply labor before the crisis is 60 percent, a number consistent with UK's employment to population and with a 1.75 percentage point decrease after the crisis.[10] Table 11 summarizes the parameter values that we used in our experiments, as well as the results.

Data in Part A of Table 11 represent average levels from the FAME sample from 2005 to 2007 with the exception of the autocorrelation coefficient and intangibles investment-to-output ratio. The autocorrelation coefficient is calculated using the entire dataset from 2005 to 2012. Intangible investment information is the average level for the UK during the relevant period. We assign the parameters in Part B for the column market "Benchmark" to make the environment as close as possible to the data. Then, the parameters in Part C are calibrated to ensure that the outcome statistics in Part A of the benchmark correspond to the statistics in the column market "Data". At the same time, the calibrated parameters in Part C of the benchmark also guarantee all the conditions of equilibrium.

As introduced in Appendix 2, the borrowing constraint is θ that governs the strength of financial frictions in the economy, which requires the debt below a fraction of its capital stock. Part C of Table 11 shows that θ decreased from 0.56 before the crisis to 0.45 after the crisis, which implies financial constraints has been tightened indeed. As a result, the TFP loss due to financial misallocation is 1.1% before the recession. After the financial crisis, the TFP loss due to financial friction increases from 1.1 to 2.1%. MX model indicate that the TFP loss due to financial frictions doubled after the crisis. However, this increase in loss is economically insignificant (1% increase), which again supports the stable overall productivity loss derived from SFA/OP/ACF/PM models after the recession.

8 Conclusions

Our paper has estimated the TFP loss in the British Isles both before and after the Great Recession using SFA/OP/ACF/PM models. The overall TFP loss ratio is around 45–51%, which is robust in all three methods. Moreover, TFP loss does not change significantly after the financial crisis. These findings imply that UK companies took some actions to successfully prevent more TFP loss after the recession. Pessoa and Van Reenen (2014) find that UK GDP per worker fell by almost 4% in the five years after 2008 was mainly due to capital shallowing, rather than fall in TFP. They claim that TFP has barely fallen at all, which is consistent with our findings. Since TFP determines long-run economic growth (Pessoa and Van Reenen 2014), the strong resilience of TFP implies there is no permanent structural change in underlying potential output growth after the recession.

However, the TFP loss due to financial friction estimated by MX model doubled after the recession, which shows the negative effect of financial constraints on productivity. Consequently, the ratio of finance-caused TFP loss increases significantly as a part of the overall TFP loss. In summary, the financial crisis resulted in stricter borrowing

[9] The probability of employed workers remaining employed.

[10] Data are from Federal Reserve Bank of St. Louis at https://fred.stlouisfed.org/series/GBREPRNA.

Table 10 Calibration and result of the Pakes–McGuire model

		Pre-financial crisis 2005–2007			Post-financial crisis 2008–2012		
		Data	MPNE	Social planner	Data	MPNE	Social planner
A. Used to calibrate model							
Average lifespan of firms		17.4	16.8		16.1	16.3	
Average investment-to-profit ratio		0.582	0.597		0.557	0.538	
B. Assigned parameters							
Constant used in investment fn.	α	3					
Cost for a GBP investment	c	1					
Maximum Number of firms	N	3					
Highest efficiency level attainable	$\overline{w}$	19					
Efficiency level for entrants	W_E	9					
Sunk cost of entry	X_E	0.2					
Lowest sunk entry cost	X_EL	0.15					
Highest sunk entry cost	X_EH	0.25					
Discount factor	β	0.92					
Prob. of outside alternative rising	δ	0.7					
Scrap value at exit	ϕ	0.1					
C. Calibrated parameters							
Vertical intercept of demand	D		4		4.95		
Fixed cost of production	f		0.3		0.8		
Capital-to-cost parameter	γ		1		1.1		
D. Result							
Average efficiency level			6.85	8.71		7.03	8.9
Max. efficiency level appeared			13	–		13	–
Average TE			0.527			0.541	
TFP Loss Ratio			0.473			0.459	

constraints that inevitably caused more TFP loss. However, UK manufacturing firms were able to reduce the loss from factors other than financial frictions (e.g., divestment and job cuts) to compensate the loss due to financial friction and hence maintain the overall TFP loss in order to survive during the recession.

Our study provides an example of how to use both non-structural and structural models to estimate overall productivity performance. Our paper can hopefully serve as a template for future studies that rely on various approaches to model TFP growth, both highly structured models and more robust but possibly less insightful non-structural alternatives. Future studies can investigate how these different non-structural and structural models perform using other database or Monte Carlo Simulations.

Acknowledgements Dr. Binlei Gong acknowledges the support from the National Natural Science Foundation of China (71903172), the Research Program for Humanities and Social Science Grant by the Chinese Ministry of Education (18YJC790034), the Soft Science Research Program of Zhejiang Province (2020C25020), and Qianjiang Talent Program.

Compliance with ethical standards

Conflict of interest The authors declare that they have no conflict of interest.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

9 Appendix 1 Pakes–McGuire model

This method computes the MPN equilibria (Maskin and Tirole 1988a, 1988b) generated under the constraints of Ericson and Pakes (1992). We employ the same notations and equations from the original work (see Pakes et al. 1993, pp. 7–36) in this section.

9.1 Model

The definition of "industry structure" is the range of efficiency levels among the various businesses in the model. A

Table 11 Calibration and results of the Midrigan–Xu model

		Pre-financial crisis 2005–2007		Post-financial crisis 2008–2012	
		Data	Benchmark	Data	Benchmark
A. Used to calibrate model					
S.D. output growth		0.3	0.3	0.3	0.3
S.D. output		1.2	1.3	1.2	1.3
1-year autocorrelation		0.98	0.97	0.98	0.98
3-year autocorrelation		0.95	0.93	0.95	0.96
5-year autocorrelation		0.93	0.90	0.93	0.94
Intangibles investment-to-output, %		12.0	8.5	12.0	3.3
Output growth rate, %		3.9	3.9	1.3	1.3
Debt-to-output		1.9	2.1	2.1	2.1
Equity-to-output		1.1	1.0	1.3	1.2
B. Assigned parameters					
Labor elasticity	α		0.6		0.6
Span of control	η		0.85		0.85
Capital depreciation	δ		8.0		8.0
Discount factor	$\beta(1+\mu)^{-1}$		0.92		0.92
Growth rate	γ		1.039		1.013
Persistence unit worker state	λ_1		0.667		0.646
Persistence zero worker state	λ_0		0.5		0.5
Relative efficiency in modern sector	$(1+\eta)\phi$		0.2		0.2
C. Calibrated parameters					
Collateral constraint	θ		0.56		0.45
Equity issuance constraint	χ		0.705		0.56
Standard deviation transitory shocks	σ_ε		0.37		0.485
Persistence transitory shocks	ρ		0.86		0.745
Cost of entering modern sector	κ		16		9
Wage	W		0.94		0.94
D. Result					
Loss misallocation, %			**1.1**		**2.1**

firm's profits arise from the industry structure as well as from the individual level of operational efficiency. Over the operating life of the firm, the efficiency of the business will change based on the stochastic environment of its operations. Decisions about investments, entries, and exits are made to achieve the highest level of future cash flow, based on expected discounted value (EDV), according to the current information set. Each business bases its decisions upon its prediction of the industry structure in the future conditional on its current information set. As a consequence, a given perception by each firm determines next period's true distribution of the industry structure. Equilibrium is reached when the projected industry structure (i.e., the predicted distribution of the efficiencies of opponents) is, in fact, the distribution that results from the predictions. Therefore, the perceived distribution of future industry structures is the same as the actual distribution that results according to the behaviors of all opponents when those behaviors are selected to achieve EDV maximization of future net cash flows generated from those perceptions. Thus, as the authors claim (Pakes and McGuire 1994, p.557), this is a Markov-perfect Nash equilibrium where the variables of choice are the investment volume combined with the exit and entry decisions.

Industry structure: $W = \{0, \ldots, \overline{w}\}$ is the set of efficiency values for each firm, where 0 is zero efficiency and $\overline{w}$ is the maximum level of efficiency. N is the maximum number of firms that can be simultaneously active in the industry. A state $[w, n]$ consists of a $w \in W^*$, $n \in N$, where $W^* = \{(w_1, \ldots, w_N) | w_j \in W,\ w_1 \geq w_2 \geq \cdots \geq w_N\}$. For any firm, w represents the economic environment, including the efficiency level of all firms, while n indicates which element of this vector is the efficiency of its own. W^* guarantees that the industry structure is represented as a weakly decreasing

N-tuple to avoid having multiple w for the same industry structure.

Investment: A firm's efficiency level for the next period is determined by a Markov process that depends on its current efficiency level, current investment, and exogenous factors. This model denotes x as the current investment, $k \in W$ as the current efficiency level of a firm, and $k' \in W$ as its efficiency level in the next period. Let τ be the effect of firm-specific investment and v the effect of other firm-invariant exogenous variables that are the same for all firms. α is a coefficient relates to self-investment rising and δ is the probability of outside alternative rising, both decide the next period efficiency level. Then the controlled Markov process for the evolution of k' is

$$k' = k + \tau - v, \tag{6}$$

where $p(\tau) = \begin{cases} (\alpha x)/(1+\alpha x), & \text{if } \tau = 1 \\ 1/(1+\alpha x), & \text{if } \tau = 0 \end{cases}$ and $p(v) = \begin{cases} \delta, & \text{if } v = 1 \\ 1-\delta, & \text{if } v = 0. \end{cases}$

Maximum level of efficiency. Equation (6) ensures that efficiency can only improve with investment and shows that the incremental efficiency in a period is bounded with probability one. Hence, there exists an upper bound of efficiency $\overline{w}$. PMM computes $\overline{w}$ as the maximum efficiency level that a monopolist would ever reach by starting with a very large efficiency level and computing the monopolist's problem to see where the monopolist stops his or her investment.

Exit and entry: Firms make the decision to exit if the future cash flow value drops below the stated scrap value of the business, which is denoted ϕ. The exiting business will only receive the current scrap value, not the current period profit. Business enters when the potential and expected future cash flow value is greater than the one-time cost of entry. The sunk cost of entry, X_E, is a random variable uniformly distributed between X_EL (lowest) and X_EH (highest). The potential entrants know the draw of X_E once they decide to enter. If they enter, they will not receive profit for that period, and enter with efficiency level W_E or W_E-1, depending on the value of v.

9.2 The algorithm

Matrices: Profit, Π, is an iteration-invariant matrix that calculates the one-shot game profit for each possible industry structure. Each iteration begins with the investment matrix, X, and the value function, V, from the output of the last iteration.

Iterative procedure: The algorithm calculates Π and iterates on X and V until the maximum of the element-by-element difference between successive iterations in these two matrices is below a specified tolerance level. For each iteration, the calculation is done separately for each of the industry structures, using the previous values of X and V. Beginning with the most efficient firm within the industry structure, its choice is updated using the most recent value of the iteration. The decision is renewed based on the value of the investment, exit, and entry. The value function is not included. These figures are used to calculate the policies for the firm with the next-highest efficiency. In turn, these updated choices are applied to the firm ranked third in efficiency, and so on.

Updating exit and entry: By comparing the value function of an incumbent competitor with the scrap value, a firm can predict if that incumbent competitor exits. The PMM defines the strategy set so that a firm must exit if it perceives that a competitor with a higher efficiency level than its own has exited. For any $[w, n]$, this model defines w' as the industry structure that results after exit has been accounted for, and m as the number of active firms in w'. After the decision of exit, this model iterates on whether there will be an entry if $m < N$. The value of future cash flows in state $[w', n]$ is compared with the one-time sunk cost in this process.

Updating investment: Each firm chooses an optimal investment policy based on its perception of future competitors. The calculation is done separately for every $[w, n] \in (W^*, N)$. The value function at the ith iteration is

$$V^i(w,n) = \max\left\{ \phi,\ sup_{x\geq 0} \begin{bmatrix} \Pi(w',n) - cx \\ + \dfrac{\beta\alpha x}{1+\alpha x} Cl(w'+e(n),n) \\ + \dfrac{\beta}{1+\alpha x} Cl(w',n) \end{bmatrix} \right\}, \tag{7}$$

where

$$Cl(w',n) = \lambda(w',n)\left\{ \sum_{\tau_1=0}^{1} \dots \sum_{\tau_n=0}^{1} \dots \sum_{\tau_N=0}^{1} \sum_{V=0}^{1} V^{i-1}[w' + W\_Ee(n_e) + \tau - iv, n] \right. \\ \left. \Pr[\tau_1|x_1^{i-1}, v] \dots \Pr[\tau_n|\ddot{x}, v] \dots \Pr[\tau_N|x_N^{i-1}, v] p(v) \right\} \\ + [1 - \lambda(w',n)]\left\{ \sum_{\tau_1=0}^{1} \dots \sum_{\tau_n=0}^{1} \dots \sum_{\tau_N=0}^{1} \sum_{V=0}^{1} V^{i-1}[w' + \tau - iv, n] \right. \\ \left. \Pr[\tau_1|x_N^{i-1}, v] \dots \Pr[\tau_n|\ddot{x}, v] \dots \Pr[\tau_N|x_N^{i-1}, v] p(v) \right\}.$$

In this value function, w' is the incumbent's efficiency level after updating for exit; $m(w')$ is the number of active competitors at $w = w'$; c is the cost in dollars of a dollar's worth of investment (equals 1 if no tax); $\lambda(w', n)$ is the probability of entry; $e(j)$ is a vector, all of whose elements are zero except for the jth element, which is one; i is a vector, all of whose elements are one; τ is the vector containing the random τ of competitors; n_e is the position of the entrant for any industry structure; $n_e = m(w') + 1$ unless the permutation cycle has been reordered; $w'_1, \ldots, w'_N$ are the elements of the vector w'; and $x_1, \ldots, x_N$ (except x_n) is the investment of the $N-1$ competitors at w'. A symbol $(\ddot{\ })$ in a summation indicates that the element is omitted. $Cl(\cdot)$ sums over the probability-weighted average of the possible states of future competitors, but not over the investing firm's own future states. It also indicates the firm's expected discounted value for each of the two possible realizations of the firm's own investment process, τ.

We denote $x^i[w', n]$ as the investment level that solves Eq. (7). To calculate it, the model first derives the optimal level of investment, $x[w', n]$, given that this investment is nonzero and that the firm does not exit. The actual level of investment, therefore, is either this number or zero, where zero is the solution if the optimal investment still leads to an exit decision or if $x[w', n]$ is negative. Let D_x denote the derivative with respect to x. The first-stage investment is thus

$$x[w',n] = \mathrm{argsolv}_x\left\{c = \beta\left[\begin{array}{c} D_x\left\{\frac{\alpha x}{1+\alpha x}\right\}Cl(w'+e(n),n) \\ -D_x\left\{\frac{\alpha x}{1+\alpha x}\right\}Cl(w',n)\end{array}\right]\right\} \tag{8}$$

It is worth noting that $D_x\left\{\frac{1}{1+\alpha x}\right\} = \frac{\alpha}{(1+\alpha x)^2} = \alpha(1-p(x))^2$, where $p(x) = \frac{\alpha x}{1+\alpha x}$. So, if $\mathrm{v1} = Cl(w'+e(n), n)$ and $\mathrm{v2} = Cl(w', n)$, the investment can be rewritten as

$$x = \mathrm{argsolv}_x\{c = \beta\alpha[1-p(x)]^2(\mathrm{v1}-\mathrm{v2})\} \Rightarrow p(x) = 1 - \sqrt{\frac{1}{\beta\alpha(\mathrm{v1}-\mathrm{v2})}}.$$

Taking the inverse of $p(x)$, it can be seen that

$$x[w',n] = \frac{p(x)}{\alpha - \alpha p(x)}.$$

It is straightforward to derive the optimal value function by plugging the optimal investment into Eq. (8) and computing

$$V^i(w,n) = \max\left\{\left[\begin{array}{c}\phi,\ \Pi(w',n) - cx[w',n] + \frac{\beta\alpha x[w',n]}{1+\alpha x[w',n]} \\ Cl(w'+e(n),n) + \frac{\beta}{1+\alpha x[w',n]}Cl(w',n)\end{array}\right]\right\}$$

If $V^i(w, n) = \phi$, then this model sets $x = 0$ with probability one. Hence, the actual investment is determined as

$$x^i[w', n] = I\{V^i(w,n) > \phi\}x[w',n],$$

where $I\{\cdot\}$ is the indicator function that takes the value of one if the condition is true, and the value of zero otherwise.

Calculating the probability of entry: After the exit decision is made, the value of entry is the value of an incumbent who realized that (1) there would be no other entry, and (2) if she enters, there would be no profits or investments in the current period for her. The expected discounted value of entering is

$$V^e(w') = \beta Cl[w' + W\_Ee[m(w')+1], m(w')+1;\ \lambda = 0].$$

A firm would like to enter if and only if $V^e > X\_E$, which is the random entry cost. Since the random cost is uniformly distributed between X_EL and X_EH, the probability of entry by an incumbent whose competitors are specified by w' and $m(w') < N$ is

$$\lambda(w',n) = \min\left\{\max\left[\frac{V^e(w') - X\_EL}{X\_EH - X\_EL}, 0\right], 1\right\}.$$

Updating N: This model starts with the one-firm problem and solves for its value function and optimal policies. Then it proceeds to the two-firm problem, using the fixed values that is solved for in the one-firm problem as the starting values for X and V:

$$V^0[(w_1,w_2),1] = V^\infty(w_1), \quad \forall w_1, w_2 \in W,$$

$$V^0[(w_1,w_2),2] = V^\infty(w_2), \quad \forall w_1, w_2 \in W,$$

where $V^\infty(\cdot)$ is the fixed point for the one-firm problem. Analogously, for the N-firm problem with $N > 2$, the starting values are

$$V^0[w,n] = \begin{cases} V^\infty[(w_1, \ldots, w_{N-1}), n], & \text{if } n < N \\ V^\infty[(w_1, \ldots, w_{N-2}, w_n), n-1], & \text{if } n = N.\end{cases}$$

The elements of X are updated in the same way as V. This process is then repeated until $\lambda(w', n) = 0$ for all (w', n) with $m(w') \geq N-1$.

9.3 Profit function

The one-shot profit function that is utilized in this model is based on a homogenous product, Nash-in-quantities (Cournot) market[11] where differences in efficiencies among firms are reflected by differences in marginal costs. Let producers' different but constant marginal costs, $\theta(w_n)$, be a firm's specific efficiency index multiplied by a common factor price index. Accordingly, if $s\tau$ and sv are the logarithms of the firm's efficiency index and of the factor price

[11] Nash-in-quantities means the quantity is the strategic variable. In other words, the strategy space for each firm contains all the finite and non-negative levels of output. And each firm chooses the output to maximize profit, taking the output choice of its opponents as fixed. Nash-in-price is a similar game, but here it is supposed that the firms choose price instead of the output.

index, respectively, then $w_n \equiv s\tau - sv$ and $\theta(w_n) = \gamma\exp(-W_n)$.

Let q_n be firm n's output, $Q = \Sigma q_n$, f be the fixed cost of production, and D be the vertical intercept of the demand curve. The profits are given by

$$\pi_n = p(Q)q_n - \theta(w_n)q_n - f = (D - Q - \theta(w_n))q_n - f.$$

The unique Nash equilibrium for this problem has quantities and price as

$$q^*_{w_n} = \max\{0, p^* - \theta_i\},\ \text{and}\ p^* = \frac{1}{n^* + 1}\left[D + \sum_{j=1}^{n^*}\theta(w_j)\right].$$

where n^* is the number of firms with positive q^*. Finally, the profit of the current period is

$$\pi(w, n) = \max\{-f, [\mathrm{p}^*(w, n) - \theta(w_n)]^2 - f\}$$

$$= \max\left\{-f, \left[\frac{1}{n^* + 1}\left[D + \sum_{j=1}^{N}\theta(w_j) - \theta(w_n)\right]^2 - f\right\}.$$

9.4 Social planner's problem

We are interested in finding out how the social planner will respond when the technology is exactly the same as that faced by the MPN competitors. The result enables us to compare the average efficiency level (and TFP loss ratio) to the MPN case.

In the homogenous products, Nash-in-quantities profits model, the social planner will set $p = \mathrm{mc}$ for the firm with the lowest marginal cost (with the highest efficiency level), and produce until supply equals demand. The methodology used to solve the social planner's issues is similar to the one used in the case of MPN competitors. The social planner's role is to maximize the stochastic profit function or the projected value of social surplus, i.e. the producer plus consumer surplus. The value function is then constructed using the value function, as defined in Eq. (7).

Since the social planner controls the entire economy, any industry structure results in only one state, not in N states. Moreover, one does not need to form perceptions about entry and exit or the behavior of cohorts, as the single agent (i.e., the social planner) controls all active firms at any time.

10 Appendix 2 Midrigan–Xu model

We also apply the benchmark model introduced by Midrigan and Xu (2014). We outline below its set-up, decision rules, definition of equilibrium, TFP function, and first-best allocation of the economy. Our discussion uses the same notations and equations from Midrigan and Xu (2014).

10.1 Set up

The economy is populated by a measure N_t of producers and a measure one of workers. The labor productivity and producer's population grow at constant rates. Producers operate either in a traditional sector that uses only labor and an unproductive technology, or in a modern sector that uses capital and labor and a more productive technology. We will focus on financial misallocation in the modern sector. A one-time sunk entry cost is required for producers in the traditional sector who want to enter into the modern sector. Moreover, one-period noncontingent security and equity claims to producers' profits are the only two kinds of financial instruments in the model.

Traditional sector producers: A certain amount, $(\gamma - 1) N_t$, of new producers enter the economy at the end of period t, but only in the traditional sector. Producers in this sector face decreasing returns on technology ($\eta < 1$) that produces output Y_t using labor L_t as the only factor of production:

$$Y_t = \exp(z + e_t)^{1-\eta}L_t^\eta \tag{9}$$

The model assumes that entrants draw the permanent productivity component z from some distribution $G(z)$, whose mean is normalized to unity. e_t is a transitory productivity component that evolves over time according to a finite-state Markov process of $\mathrm{E} = (e_1, \ldots, e_T)$ with transition probabilities $f_{i,j} = \Pr(e_{t+1} = e_j | e_t = e_i)$. Entrants draw their initial productivity component e_i from the stationary distribution associated with f, which we denote with $\bar{f}_i$.

All producers in the traditional sector aim to maximize their lifetime utility, which is $E_0 \sum_{t=0}^{\infty} \beta^t \log(C_t)$. However, the budget constraints they face depend on whether remaining in the traditional sector or switching to the modern sector.

On the one hand, the budget constraint for those who stay in the traditional sector is

$$C_t = Y_t - WL_t - (1 + r)D_t + D_{t+1}, \tag{10}$$

where D_t denote the producer's debt position, which is nonpositive since these producers are not allowed to borrow. All entering producers have no wealth, i.e. the initial D is equal to zero. Moreover, W and r are the equilibrium wage and interest rate.

On the other hand, traditional sector producers who enter the modern sector require an investment equal to $\exp(z)\kappa$ units of output, which is proportional to the permanent productivity component. Besides internal funds, both of the two financial instruments including one-period risk-free debt and equity claims to future profits are potential channels to finance the physical capital, K_{t+1}, and intangible capital, $\exp(z)\kappa$. In terms of debt, the borrowing constraint is

$$D_{t+1} \leq \theta(K_{t+1} + \exp(z)\kappa). \tag{11}$$

where $\theta \in [0, 1]$ governs the strength of financial frictions in the economy, which requires the debt below a fraction of its capital stock. For equity claims, denote P_t as the price of the claim to the entire stream of profits, where profits are defined as $\Pi_t^m = Y_t - WL_t - (r + \delta)K_t$, where δ is the capital depreciation rate. The model assumes that producers can only issue claims to a fraction, $\theta\chi$, of their future profits, where $\chi \in [0, 1]$. θ is characterizing the degree of financial development of the economy since it decides the producer's ability to both borrow and issue equity. The budget constraint of a producer that enters the modern sector is therefore

$$C_t + K_{t+1} + \exp(z)\kappa = Y_t - WL_t - (1 + r)D_t + D_{t+1} + \theta\chi P_t. \tag{12}$$

Modern sector producers: The production function for the producers in the modern sector is

$$Y_t = \exp(z + e_t + \phi)^{1-\eta}\left(L_t^\alpha K_t^{1-\alpha}\right)$$

where $\phi \geq 0$ determines the relative productivity of this sector, α controls the share of labor in production, and K_t is the amount of capital used in the previous period.

Producers in the modern sector can save and borrow at the risk-free rate, r, subject to the constraint (11). Their budget constraint is

$$C_t + K_{t+1} - (1 - \delta)K_t = Y_t - WL_t - (1 + r)D_t - \theta\chi\Pi_t^m + D_{t+1}. \tag{13}$$

The model assumes, as is standard in the investment literature, that output at date $t + 1$ is produced with capital held in period t. The choice of how much to invest at the end of period t is, however, measurable with respect to e_{t+1}. This assumption of timing explains why the expected return to equity equals the risk free return.

Workers: A unit measure of workers is available in the economy, each of whom supplies $\gamma^t\nu_t$ efficiency units of labor, where ν_t is the worker's idiosyncratic efficiency that evolves over time according to a finite-state Markov process. These workers have the same log preferences (utility function) as producers do. However, their budget constraint is

$$c + a_{t+1} + \int P_t^i \omega_{t+1}^i di = W\gamma^t\nu_t + (1 + r)a_t + \int \left(P_t^i + \Pi_t^{m,i}\right)\omega_t^i di,$$

where a_t denote a worker's holdings of risk-free assets and ω_t^i denote the number of shares he or she owns of producer i. The total asset holdings, $a_{t+1} + \int P_t^i \omega_{t+1}^i di$, are non-negative because the model assumes that workers cannot borrow.

Once again, there is no aggregate risk in this economy due to the assumption of timing. As a result, the lack of arbitrage implies that the return on the risk-free security is equal to the expected return on equity claims:

$$(1 + r) = \frac{E_t\left[P_{t+1}^i + \Pi_{t+1}^{m,i}\right]}{P_t^i}.$$

10.2 Recursive formulation and decision rules

Modern sector producers: The risk-free assumption on capital implies that producer profits are solely a function of its net worth, which is denoted as $A = K - D$. Moreover, profits, output, and the optimal choice of capital and labor are all homogeneous of degree one in $(A, \exp(z))$ so this model can rescale all variables by $\exp(z)$ including the rescaled net worth $a = A/\exp(z)$. Given the new notation, the Bellman equation is

$$V^m(a, e_i) = \max_{a',c} \log(c) + \beta \sum_m f_{i,j} V^m(a', e_j). \tag{14}$$

Similarly, the budget constraint in Eq. (13) can be rewritten as

$$c + a' = (1 - \theta\chi)\pi^m(\alpha, e) + (1 + r)a, \tag{15}$$

where

$$\pi^m(\alpha, e) = \max_{k,l} \exp(e + \phi)^{1-\eta}\left(l^\alpha k^{1-\alpha}\right)^\eta - Wl - (r + \delta)k. \tag{16}$$

Furthermore, the borrowing constraint in Eq. (11) reduces to

$$k \leq \frac{1}{1 - \theta}a + \frac{\theta}{1 - \theta}\kappa \tag{17}$$

This model characterizes the producer's net worth accumulation decision of the producer by

$$\frac{1}{c(a, e_i)} = \beta \sum f_{i,j}\left[(1 + r) + \frac{1}{1 - \theta}\mu(a', e_j)\right]\frac{1}{c(a', e_j)}, \tag{18}$$

where $\mu(a, e)$ is the multiplier on the borrowing constraint (17). The producer's return to savings increases with the expectation that the borrowing constraint will be binding in future periods. Therefore, the producers have the incentive to accumulate net worth.

Accordingly, the decisions on the optimal level of capital and labor simplifies to

$$\alpha\eta\frac{y(a,e)}{l(a,e)}=W \tag{19}$$

and

$$(1-\alpha)\eta\frac{y(a,e)}{k(a,e)}=r+\delta+\mu(a,e). \tag{20}$$

Dispersion of net worth and productivity of businesses due to borrowing constraints causes dispersion in the marginal product of capital of individual producers. In turn, this causes TFP reductions due to misallocation. In the rescaled formulation of the problem, it is worth noting that the producer's permanent productivity component, z, has no independent effect on allocations.

Traditional sector producers: The next thing to consider is the problem of producers in the traditional sector. Since capital is not an input for these producers, their net worth is $a=-d$. This model also denotes x as their savings. The Bellman equation for such producers is

$$V^{\tau}(a,e_i)=\max_{a',c}\log(c)+\beta\max\left\{\sum_j f_{i,j}V^{\tau}(a',e_j),\ \sum_m f_{i,j}V^{m}(a',e_j)\right\},$$

subject to

$$c+x=\pi^{\tau}(e)+(1+r)a \tag{21}$$

where

$$\pi^{\tau}(e)=\max_l \exp(e)^{1-\eta}l^{\eta}-Wl.$$

In each period, the producer's decision on whether to stay in the traditional sector or switch to the modern sector depends on the relative value of these two options. This decision also determines the evolution of its net worth. A producer who remains in the traditional sector simply inherits its past savings, $a'=x$, while a producer that enters the modern sector has

$$a'=x-\kappa+\theta\chi p(a',e_i) \tag{22}$$

where $p(a', e_i)$ is the rescaled price of the equity claim to that satisfies

$$p(a,e_i)=\frac{1}{1+r}\sum_j f_{i,j}\left[p(a',e_j)+\pi^{m}(a',e_j)\right] \tag{23}$$

The producers in the modern sector may have negative net worth since they can borrow against the intangible capital. Besides the collateral constraint in Eq. (17), the natural borrowing constraint is

$$a>a_{\min}=-\frac{(1-\theta\chi)\pi^{m}(a_{\min},e_1)}{r} \tag{24}$$

which guarantees the producer's solvency even under the worst possible sequence of productivity shocks. This constraint may be more stringent than the collateral constraint and motivate producers to accumulate enough savings before entering the modern sector even in the absence of a collateral constraint.

10.3 Equilibrium

Denote $n_t^m(a,e)$ as the measure of modern-sector producers and $n_t^{\tau}(a, e)$ as the measure of traditional sector producers. The population of producers in these two sectors sum to $N_t=\gamma^t$: $\int_{A\times E}dn_t^m(a,e)+\int_{A\times E}dn_t^{\tau}(a,e)=N_t$.

On the one hand, the number of producers in the modern sector evolves according to

$$\begin{aligned}n_{t+1}^{m}(A,e_j)&=\int_A\sum_i f_{i,j}\mathrm{I}_{\{a^m(a,e_i)\in A\}}dn_t^m(a,e_i)\\&\quad+\int_A\sum_i f_{i,j}\mathrm{I}_{\{\xi(a,e_i)\in A\}}dn_t^{\tau}(a,e_i),\end{aligned} \tag{25}$$

where $\xi(a, e)$ is an indicator of whether a producer in the traditional sector switches, $A=[\underline{a},\overline{a}]$ is the compact set of values that a producer's net worth can take and A is a family of its subsets, $a^m(.)$ is the amount of net worth for a producer in the modern sector, and $a^{\tau,s}(.)$ is the savings decision of a producer who switches.

On the other hand, the measure of producers in the traditional sector is

$$\begin{aligned}n_{t+1}^{\tau}(A,e_j)&=\int_A\sum_i f_{i,j}\mathrm{I}_{\{\xi(a,e_i)=0,a^{\tau}(a,e_i)\in A\}}dn_t^{\tau}(a,e_i)\\&\quad+(\gamma-1)N_t\mathrm{I}_{\{0\in A\}}\overline{f}_J,\end{aligned} \tag{26}$$

where $\overline{f_j}$ is the stationary distribution of the transitory productivity and $a^{\tau}(.)$ is the net worth of a producer that stays in the traditional sector.

A balanced growth equilibrium must satisfy the following five conditions:

(I) the labor market clearing condition:

$$\int_{A\times E}l^{\tau}(e)dn_t^{\tau}(a,e)+\int_{A\times E}l^{m}(a,e)dn_t^{m}(a,e)=L_t=\gamma^t,$$

(II) the asset market clearing condition:

$$A^w_{t+1} + \sum_{i=m,\tau} \int_{A\times E} \sum a^i_{t+1}(a,e)dn^i_{t+1}(a,e) = \int_{A\times E} k^m_{t+1}(a,e)dn^m_{t+1}(a,e), \tag{27}$$

or

$$C_t + K_{t+1} - (1-\delta)K_t + X_t = Y_t, \tag{28}$$

(III) producer and worker optimization,
(IV) the no-arbitrage condition in Eq. (23),
(V) the laws of motion for the measures in Eqs. (25) and (26).

All variables with time subscripts grow at a constant rate γ while all other variables are fixed. Solving the balanced growth equilibrium is equal to solving the stationary system where all the time-variant variables are rescaled by γ^t.

10.4 Efficient allocations

The value of TFP in the economy is reduced by financial frictions, which occur in two ways: either by affecting a company's entry decision into the modern sector or by causing losses in the modern sector due to misallocation. The strength of these two paths is defined using two separate computations. The first computation determines the level of TFP losses in the modern sector as a result of capital misallocation when the number of modern producers (n^m) is given. The equilibrium of the model is taken as the stationary level. This calculation is similar to the one that was stated by Hsieh and Klenow (2009). The second computation calculates the optimal allocation of producers across the traditional and modern sectors by solving a planner's problem (i.e. n^m is not fixed). The broader question in this calculation is identifying the level of consumption in the economy and how it is limited by financial frictions that develop along the way at both intensive and extensive margins.

TFP losses from misallocation in the modern sector: Let i index producers and M be the set of all producers in the modern sector. Also, let L and K be the total amount of labor and capital used in that sector, respectively. Integrating the decision rules (19) and (20) across producers, the total amount of output produced by the modern sector is

$$Y = \underbrace{\exp(\phi)\frac{\left(\int_{i\in M}\exp(e_i)(r+\delta+\mu_i)^{-\frac{(1-\alpha)\eta}{1-\eta}}di\right)^{1-\alpha\eta}}{\left(\int_{i\in M}\exp(e_i)(r+\delta+\mu_i)^{-\frac{\alpha\eta-1}{1-\eta}}di\right)^{(1-\alpha)\eta}}}_{=TFP}(L^\alpha_t K^{1-\alpha}_t)^\eta \tag{29}$$

This expression shows that TFP of the modern sector is determined by the exogenous productivity gap, ϕ, and an endogenous component that depends on the measure of producers, their efficiency, and the extent to which they are bind.

To calculate the efficient level of TFP given a measure of M producers, the model allocates capital and labor across producers so that the marginal products of capital and labor are the same across producers in order to maximize total output in the modern sector. Accordingly, the efficient level of output is given by

$$Y^e = \underbrace{\exp(\phi)^{1-\eta}\left(\int_{i\in M}\exp(e_i)di\right)^{1-\eta}}_{=TFP^e}(L^\alpha K^{1-\alpha})^\eta. \tag{30}$$

Comparing Eqs. (29) and (30) and using the fact that the shadow cost of capital, $r+\delta+\mu$, is proportional to its average product, as in Eq. (20), the TFP losses from misallocation are

$$\text{TFP losses} = \log\left(\int_{i\in M}\exp(e_i)\right)^{1-\eta} - \log\frac{\left(\int_{i\in M}\exp(e_i)\left(\frac{y_i}{k_i}\right)^{-\frac{(1-\alpha)\eta}{1-\eta}}\right)^{1-\alpha\eta}}{\left(\int_{i\in M}\exp(e_i)\left(\frac{y_i}{k_i}\right)^{\frac{\alpha\eta-1}{1-\eta}}\right)^{(1-\alpha)\eta}}. \tag{31}$$

To clarify Eq. (31), suppose that the logarithm of y_i/k_i and e_i are jointly normally distributed. Equation (31) then reduces to

$$\text{TFP losses} = \frac{1}{2}\frac{(1-\alpha\eta)(1-\alpha)\eta}{1-\eta}\text{var}(\log(y_i/k_i)), \tag{32}$$

so that the TFP losses are proportional to the variance of the average product of capital. In other words, higher variability in the average product of capital across producers generates more TFP losses.

Efficient (first-best) allocations: To calculate the efficient allocation, this model must also derive the optimal number of producers across the two sectors. This can be done by solving the social planner's problem that is only constrained by the aggregate resource constraint in Eq. (28) and by the production technologies that we have assumed. Accordingly, their study chooses the amount of capital, K, the number of producers in the two sectors, n^τ_i and n^m_i, and the allocation of labor across those sectors, L^τ and L^m, to maximize

$$\underbrace{\left(\sum\nolimits_i \exp(e_i)n^\tau_i\right)^{1-\eta}(L^\tau)^\eta}_{\text{output in traditional sector}} + \underbrace{\left(\sum\nolimits_i \exp(e_i+\phi)n^m_i\right)^{1-\eta}\left((L^m)^\alpha(K)^{1-\alpha}\right)^\eta}_{\text{output in modern sector}} - \underbrace{\left(\delta+\frac{\gamma}{\beta}-1\right)K}_{\text{cost of capital}} - \underbrace{\frac{(\gamma-1)\kappa\sum_i n^m_i}{\beta}}_{\text{sunk cost of entering}}, \tag{33}$$

subject to the restrictions on the measurements implied by Markov transition probabilities, $f_{i,j}$, and to the labor constraint, $L^{\tau} + L^{m} = 1$.

10.5 Summary

In short, the model has three kinds of players: workers, traditional producers, and modern producers. Traditional producers use only labor and unproductive technology, and cannot borrow money. Modern producers, on the other hand, use capital, labor, and more productive technology; they can also borrow. Traditional producers can become modern producers, but to do so they must incur a sunken entry fee, and they are allowed to borrow and issue claims to part of the future profit during that period of transformation. The amount that a producer can borrow is subject to collateral constraints. Workers face uninsurable idiosyncratic labor income risk and have access to financial markets. There are two types of financial instruments available: a one-period non-contingent security and equity claims to producers' profits.

These three kinds of players all try to maximize their lifetime utility. The equilibrium requires (I) a labor market clearing condition, (II) an asset market clearing condition, (III) producer and worker optimization, (IV) the no-arbitrage condition, and (V) the laws of motion. Equation (3) is used to calculate the TFP loss due to financial misallocation. On the one hand, the actual TFP level in the equilibrium can be derived under this setup. On the other hand, the efficient level of TFP, TFP^{e}, is the solution to the planner's problem that is not restricted in any way concerning the allocation of labor and capital across firms.

References

Ackerberg DA, Caves K, Frazer G (2015) Identification properties of recent production function estimators. Econometrica 83:2411–2451
Agrell PJ, Bogetoft P (2017) Regulatory benchmarking: models, analyses and applications. Data Envelopment Anal J 3:49–91
Agrell PJ, Bogetoft P (2018) Theory, techniques, and applications of regulatory benchmarking and productivity analysis, the Oxford handbook of productivity analysis, Oxford University Press, New York
Aigner D, Lovell CA, Schmidt P (1977) Formulation and estimation of stochastic frontier production function models. J Econ 6:21–37
Amsler C, Prokhorov A, Schmidt P (2016) Endogeneity in stochastic frontier models. J Econ 190:280–288
Arellano M, Bond S (1991) Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. Rev Economic Stud 58:277–297
Asmild M, Baležentis T, Hougaard JL (2016) Multi-directional productivity change: MEA-Malmquist. J Product Anal 46:109–119
Bai J (2009) Panel data models with interactive fixed effects. Econometrica 77:1229–1279
Bai J, Ng S (2002) Determining the number of factors in approximate factor models. Econometrica 70:191–221
Ball VE, San-Juan-Mesonada C, Ulloa CA (2014) State productivity growth in agriculture: catching-up and the business cycle. J Product Anal 42:327–338
Basu S, Fernald JG (1997) Returns to scale in US production: estimates and implications. J Political Econ 105:249–283
Battese GE, Coelli TJ (1992) Frontier production functions, technical efficiency and panel data: with application to paddy foarmers in India. J Productivity Anal 3:153–169
Benfratello L, Sembenelli A (2002) Research joint ventures and firm level performance. Res Policy 31:493–507
Blundell R, Bond S (1998) Initial conditions and moment restrictions in dynamic panel data models. J Econ 87:115–143
Bogetoft P (2013) Performance benchmarking: measuring and managing performance. Springer Science & Business Media, New York, USA
Boucinha M, Ribeiro N, Weyman-Jones T (2013) An assessment of Portuguese banks' efficiency and productivity towards euro area participation. J Product Anal 39:177–190
Bughin J, Dimson J, Hunt V, Allas T, Krishnan M, Mischke J, et al. (2018) Solving the United Kingdom's productivity puzzle in a digital age. McKinsey Global Institute Discussion Paper
Burnside C (1996) Production function regressions, returns to scale, and externalities. J Monetary Econ 37:177–201
Chadha J, Crystal A, Pearlman J, Smith P, Wright S (2016) The UK economy in the long expansion and its aftermath. Cambridge University Press, Cambridge, UK
Chamberlain G (1984) Panel data. Handb Econ 2:1247–1318
Comin D (2010) Total factor productivity. In: Economic growth. Palgrave Macmillan, London, UK, 260–263
Cornwell C, Schmidt P, Sickles RC (1990) Production frontiers with cross-sectional and time-series variation in efficiency levels. J Econ 46:185–200
Cummins JD, Xie X (2013) Efficiency, productivity, and scale economies in the US property-liability insurance industry. J Product Anal 39:141–164
Curi C, Lozano-Vivas A (2015) Financial center productivity and innovation prior to and during the financial crisis. J Product Anal 43:351–365
Derbali A (2016) Global systemic risk ranking of Chinese financial institutions. J Chin Gov 1:340–372
Draca M, Machin S, Van Reenen J (2011) Minimum wages and firm profitability. American Economic. J Appl Econ 3:129–151
Eberhardt M, Helmers C (2010) Untested assumptions and data slicing: a critical review of firm-level production function estimators. Department of Economics, University of Oxford, Oxford, UK
Ericson R, Pakes A (1992) An alternative theory of firm and industry dynamics. Cowles Foundation for Research in Economics at Yale University, New Haven, Connecticut, USA
Faggio G, Salvanes KG, Van Reenen J (2010) The evolution of inequality in productivity and wages: panel data evidence. Ind Corp Change 19:1919–1951
Forni M, Hallin M, Lippi M, Reichlin L (2000) The generalized dynamic-factor model: Identification and estimation. Rev Econ Stat 82:540–554
Del Gatto M, Di Liberto A, Petraglia C (2011) Measuring productivity. J Economic Surv 25:952–1008
Girma S, Görg H, Hanley A (2008) R&D and exporting: a comparison of British and Irish firms. Rev World Econ 144:750–773
Gong B (2018a) Agricultural reforms and production in China changes in provincial production function and productivity in 1978–2015. J Dev Econ 132:18–31

Gong B (2018b) The Shale Technical Revolution—Cheer or Fear? Impact Analysis on Efficiency in the Global Oilfield Service Market. Energy Policy 112:162–172

Gong B (2018c) Different behaviors in natural gas production between national and private oil companies: Economics-driven or environment-driven? Energy Policy 114:145–152

Gong B (2019) New Growth Accounting. American Journal of Agricultural Economics forthcoming

Gong B (2020) Effects of ownership and business portfolio on production in oil and gas industry. Energy J 41:33–54

Goodridge P, Haskel J, Wallis G (2012) UK innovation index: productivity and growth in UK industries. CEPR discussion papers, London, UK

Graham DJ, Melo PS, Jiwattanakulpaisarn P, Noland RB (2010) Testing for causality between productivity and agglomeration economies. J Regional Sci 50:935–951

Graham DJ (2007) Agglomeration, productivity and transport investment. J Transp Econ Policy 41:317–343

Greene WH (1990) A gamma-distributed stochastic frontier model. J Econ 46:141–163

Griffith R, Harrison R, Reenen JV (2006) How special is the special relationship? Using the impact of U.S. R&D spillovers on U.K. firms as a test of technology sourcing. Am Economic Rev 96:1859–1875

Guariglia A, Mateut S (2010) Inventory investment, global engagement, and financial constraints in the UK: evidence from micro data. J Macroecon 32:239–250

Harris R, Li QC (2008) Evaluating the contribution of exporting to UK productivity growth: some microeconomic evidence. World Econ 31:212–235

Hausman JA, Taylor WE (1981) Panel data and unobservable individual effects. Econometrica 49:1377–1398

Helmers C, Rogers M (2010) Innovation and the survival of new firms in the UK. Rev Ind Organ 36:227–248

Helmers C, Rogers M (2011) Does patenting help high-tech start-ups? Res Policy 40:1016–1027

Helmers C, Rogers M, Schautschick P (2011) Intellectual property at the firm-level in the UK: the Oxford firm-level intellectual property database. Department of Economics, University of Oxford, Oxford, UK

Hsieh C-T, Klenow PJ (2009) Misallocation and manufacturing TFP in China and India. Q J Econ 124:1403–1448

Kneip A, Sickles R, Song W (2003) On estimating a mixed effects model with applications to the U.S. banking industry. Rice University, Mimeo, Houston, Texas, USA

Kneip A (1994) Nonparametric estimation of common regressors for similar curve data. Ann Stat 22:1386–1427

Kumbhakar SC, Wang H-J (2005) Estimation of growth convergence using a stochastic production frontier approach. Econ Lett 88:300–305

Kutlu L (2018) A distribution-free stochastic frontier model with endogenous regressors. Econ Lett 163:152–154

Maskin E, Tirole J (1988a) A theory of dynamic oligopoly, I: overview and quantity competition with large fixed costs. Econometrica 56:549–569

Maskin E, Tirole J (1988b) A theory of dynamic oligopoly, II: price competition, kinked demand curves, and Edgeworth cycles. Econometrica 56:571–599

McGuckin RH, Nguyen SV, Taylor JR, Waite CA (1992) Post-reform productivity performance and sources of growth in Chinese industry: 1980–85. Rev Income Wealth 38:249–266

Meeusen W, Van den Broeck J (1977) Efficiency estimation from Cobb-Douglas production functions with composed error. Int Economic Rev 18:435–444

Midrigan V, Xu DY (2014) Finance and misallocation: evidence from plant-level data. Am Economic Rev 104:422–458

Mollisi V, Rovigatti G (2017) Theory and practice of TFP estimation: the control function approach using Stata, CEIS Working Paper, Rome, Italy

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

Olley GS, Pakes A (1996) The dynamics of productivity in the telecommunications equipment industry. Econometrica 64:1263–1297

Pakes A, Gowrisankaran G, McGuire P (1993) Implementing the Pakes-McGuire algorithm for computing Markov perfect equilibria in Gauss. Working paper. Yale University, New Haven

Pakes A (1996) Dynamic structural models, problems and prospects: mixed continuous discrete controls and market interaction, In: Advances in Econometrics: Vol. 2, Sixth World Congress, Cambridge University Press, New York, USA

Pakes A, McGuire P (1994) Computing Markov-perfect Nash equilibria: numerical implications of a dynamic differentiated product model. Rand J Econ 25:555–589

Perelman S (1995) R&D, technological progress and efficiency change in industrial activities. Rev Income Wealth 41:349–366

Pesaran MH (2006) Estimation and inference in large heterogeneous panels with a multifactor error structure. Econometrica 74:967–1012

Pessoa JP, Van Reenen J (2014) The UK productivity and jobs puzzle: does the answer lie in wage flexibility? Econ J 124:433–452

Pitt MM, Lee L-F (1981) The measurement and sources of technical inefficiency in the Indonesian weaving industry. J Dev Econ 9:43–64

Riley R, Rincon-Aznar A, Samek L (2018) Below the aggregate: a sectoral account of the UK productivity puzzle. Economics Statistics Centre of Excellence Discussion Paper, London, UK

Schmidt P, Sickles RC (1984) Production frontiers and panel data. J Bus Economic Stat 2:367–374

Sickles RC (2005) Panel estimators and the identification of firm-specific efficiency levels in parametric, semiparametric and nonparametric settings. J Econ 126:305–334

Stevenson RE (1980) Likelihood functions for generalized stochastic frontier estimation. J Econ 13:57–66

Stock JH, Watson MW (2002) Forecasting using principal components from a large number of predictors. J Am Stat Assoc 97:1167–1179

Tenreyro S (2018) The fall in productivity growth: causes and implications. Speech at Preston Lecture Theatre, Queen Mary University of London, London, UK

Timmer MP, Voskoboynikov IB (2014) Is mining fuelling long-run growth in Russia? Industry productivity growth trends since 1995. Rev Income Wealth 60:S398–S422

Van Dijk M, Szirmai A (2011) The micro-dynamics of catch-up in Indonesian paper manufacturing. Rev Income Wealth 57:61–83