# A Quantal Response Statistical Equilibrium Model of Induced Technical Change in an Interactive Factor Market: Firm-Level Evidence in the EU Economies

**Jangho Yang †**

Department of Economics, The New School for Social Research, 6 E 16th Street, New York, NY 10003, USA; yangj994@newschool.edu; Tel.: +1-347-276-4512
† Current affiliation: Oxford Martin School, University of Oxford, 34 Broad St, Oxford OX1 3BD, UK.

**Abstract:** This paper studies the pattern of technical change at the firm level by applying and extending the Quantal Response Statistical Equilibrium model (QRSE). The model assumes that a large number of cost minimizing firms decide whether to adopt a new technology based on the potential rate of cost reduction. The firm in the model is assumed to have a limited capacity to process market signals so there is a positive degree of uncertainty in adopting a new technology. The adoption decision by the firm, in turn, makes an impact on the whole market through changes in the factor-price ratio. The equilibrium distribution of the model is a unimodal probability distribution with four parameters, which is qualitatively different from the Walrasian notion of equilibrium in so far as the state of equilibrium is not a single state but a probability distribution of multiple states. This paper applies Bayesian inference to estimate the unknown parameters of the model using the firm-level data of seven advanced OECD countries over eight years and shows that the mentioned equilibrium distribution from the model can satisfactorily recover the observed pattern of technical change.

**Keywords:** induced technical change; statistical equilibrium; bounded rationality; cost minimizing behavior; quantal response; factor price

**JEL Classification:** C10; D01; D21; D22; D80; O33

## 1. Introduction

One of the most interesting features of the patterns of technical change in the modern economy is that the empirical frequency distributions of the rate of changes in labor productivity, capital productivity, and the realized rate of cost reduction (a negative total factor productivity growth) have a stable shape in many different economies over different time periods. Figure 1 displays the log frequency distribution of these variables for the UK firms from 2006–2013, whose patterns are consistently observed in other countries, as we will see in the following section:

All distributions exhibit the same highly peaked symmetric pattern, yielding a well-behaved tent-shaped distribution. This consistency of the empirical frequency distributions implies that a systematic force in the process of technical change generates what is called in thermodynamics *statistical equilibrium* [1,2]. Statistical equilibrium is defined as the most likely state of the system in the form of a probability distribution. Unlike the conventional notions of equilibrium in economics such as the market-clearing Walrasian equilibrium, whose main theoretical power is to prove the existence of the equilibrium state of the system expressed as the fixed point, statistical equilibrium inherently predicts the central tendency and the fluctuations around it simultaneously. Consequently, it provides a useful framework for analyzing the systemic force underlying the patterns of observed frequency

distributions. (There is a wide body of economics literature on statistical equilibrium applications such as [3–10]. For a survey on information theoretic approaches in economics, see [11–14].)

## Marginal Log Frequency of $\gamma$, $\chi$, $\zeta$ :UK
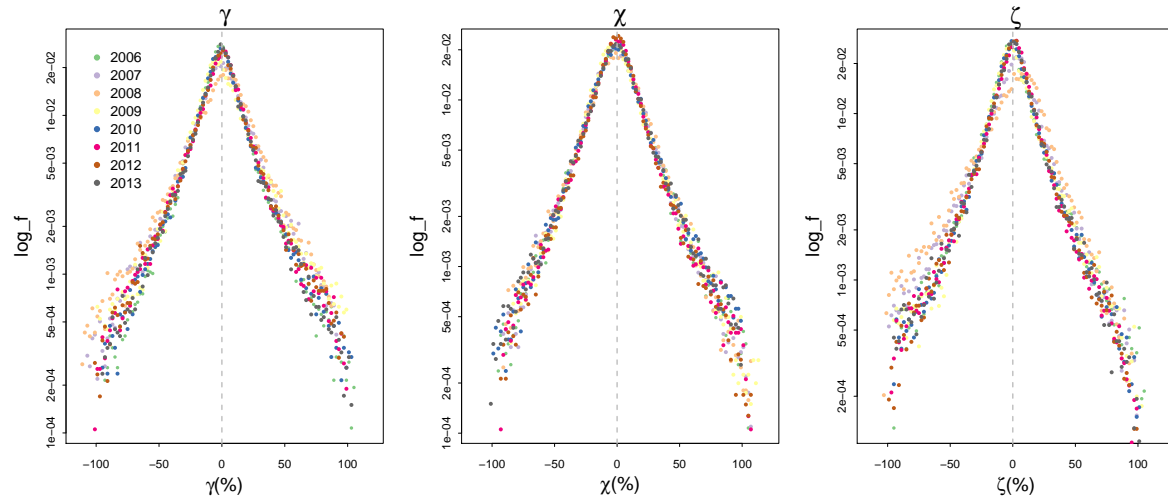


**Figure 1.** The frequency distribution of the growth rate of labor and capital productivity, $\gamma, \chi$ and the rate of cost reduction $\zeta$ of the UK from 2006–2013 with log scale on the vertical axis. The histograms are centered for each year. Both distributions exhibit a peaked tent shape. Data Source: ORIBS-AMADEUS.

This paper applies and extends the theoretical framework of the Quantal Response Statistical Equilibrium model (QRSE) recently proposed by [15], and develops a model of two systemic forces that govern the process of firms' technical progress. The paper also utilizes an entropy constrained model of induced technical change (ITC) as the baseline model of a firm's behavior in adopting a new technology. The ITC model makes a behavioral assumption that the firm tries to maximize the rate of cost reduction, and therefore, responds with a higher probability to the higher potential cost reduction the new technology would bring about. The firm, however, is assumed to have a limited capacity to process market signals so that there is a positive uncertainty in adopting an optimal technology. On the other hand, the model assumes that the adoption of a new technology by the firm makes an impact on the factor market and adversely changes the factor-price ratio of the firm against the firm's initial choice.

The equilibrium distribution of this model is a unimodal distribution of the rate of cost reductions with four parameters, qualitatively predicting the observed peaked distribution of the rate of cost reduction. We then utilize a Bayesian inference to recover the four unknown parameters of the model by fitting the derived statistical equilibrium distribution to the observed distribution of a firm's rate of cost reductions in the seven different advanced EU countries over eight years. The empirical result shows that, during the financial crisis in 2007–2009, the firms became more reluctant to adopt a new technology and needed a higher premium on the potential rate of cost reduction to do so. Also, there was a higher degree of uncertainty in adopting a new technology, so that the economic gains from the technical change were reduced.

The paper consists of four parts. Section 2 provides a description of the data we use in the paper and displays the firm-level frequency distributions of a few technological variables such as the rate of changes in labor productivity, capital productivity, and the realized rate of cost reductions. Section 3 discusses the QRSE model of technical change. The statistical equilibrium of the model will be derived in this section. Section 4 sets up a Bayesian model to estimate the unknown parameters of the statistical equilibrium distribution and recovers the posterior distributions of the rate of cost reductions and the adoption decision. Section 5 discusses further the methodological advantages of the statistical equilibrium approach.

## 2. Patterns of Technical Change

### 2.1. Data

We use the AMADEUS database where OECD firm-level data is available (I am grateful to Columbia University for the access to AMADEUS database.) and extract four variables: wages (W), the number of employees (L), total asset (K), and earnings before interest and taxes (EBIT) for 9 years of observations in 7 different advanced EU countries: the United Kingdom, Germany, France, Italy, Spain, Sweden, and Portugal. The first five countries are the largest EU economies and their combined GPD shares around 70% of the entire EU GDP. Sweden and Portugal are included to see if relatively small sized economies have the same patterns as other larger economies.

Labor productivity $x$ and capital productivity $\rho$ are defined as the total value added divided by total labor and total capital invested, respectively. Since the sum of EBIT and wages is the value added of the firm, $x$ and $\rho$ in year $t$ are calculated as $x_t = (W_t + \text{EBIT}_t)/L_t$ and $\rho_t = (W_t + \text{EBIT}_t)/K_t$. The growth rate of labor and capital productivity is obtained as $\gamma_t = (x_t - x_{t-1})/x_{t-1}$ and $\chi_t = (\rho_t - \rho_{t-1})/\rho_{t-1}$. For a detailed discussion on the accounting framework, see [16].

Assuming the constancy of the wage and profit rates, the growth rate of cost reduction can be expressed as the average of $\gamma$ and $\chi$ weighted by the unit labor and unit capital costs, $\zeta_t = \omega_t \gamma_t + \pi_t \chi_t$, where $\omega = W_t / (W_t + \text{EBIT}_t)$ and $\pi = \text{EBIT}_t / (W_t + \text{EBIT}_t)$, respectively. ($\zeta$ can be derived by taking the log derivative of the unit total cost given the wage and profit rate. The total costs $C$ is the sum of the total labor and capital cost: $C = rK + wL$, where $r = EBIT/K$ and $w = W/L$ are the rate of profit and the wage rate. Therefore, $\frac{d(C/Y)}{C/Y} = \omega d(x)/x + \pi d(\rho)/\rho = \omega\gamma + \pi\chi$.) It is instructive to note that $\zeta$ is mathematically equivalent to the negative total factor productivity growth. We exclude those firms whose value added (W + EBIT) is negative. This effectively sets the lower bound of the growth rates to −100%. For simplicity, we also confine the upper bound of the growth rate to 100%. This is one way of excluding extreme values from the noisy firm-level data. Since the focus of the paper is to analyze some salient patterns of the observed distributions, restricting the data to economically plausible range can be a good first approximation. Around 10% of the data points have been removed after this manipulation. The following is the number of observations for 7 countries for 8 years: the United Kingdom = 188,223, Germany = 86,306, France = 75,894 , Italy = 345,037, Spain = 524,705, Portugal = 181,195, Sweden = 211,237.

### 2.2. Empirical Distributions of $\gamma$, $\chi$ and, $\zeta$

We first present the log-frequency distributions of the growth rate of labor and capital productivity, $\gamma$ and $\chi$ in Figure 2. See Appendix A for a summary statistics of the distributions.

Both $\gamma$ and $\chi$ in all countries in different years show the same patterns as the ones from the UK mentioned in the introduction, exhibiting a fairly stable tent shaped distribution. This strong regularity is also observed in the frequency distribution of $\zeta$ as is shown in Figure 3.

The highly peaked symmetric distribution of $\zeta$ consistently observed in many different countries over a decade implies that there is a systematic force in the process of technical change that governs the deviation of the rate of cost reduction from its central tendency, which we will explore in the following section.
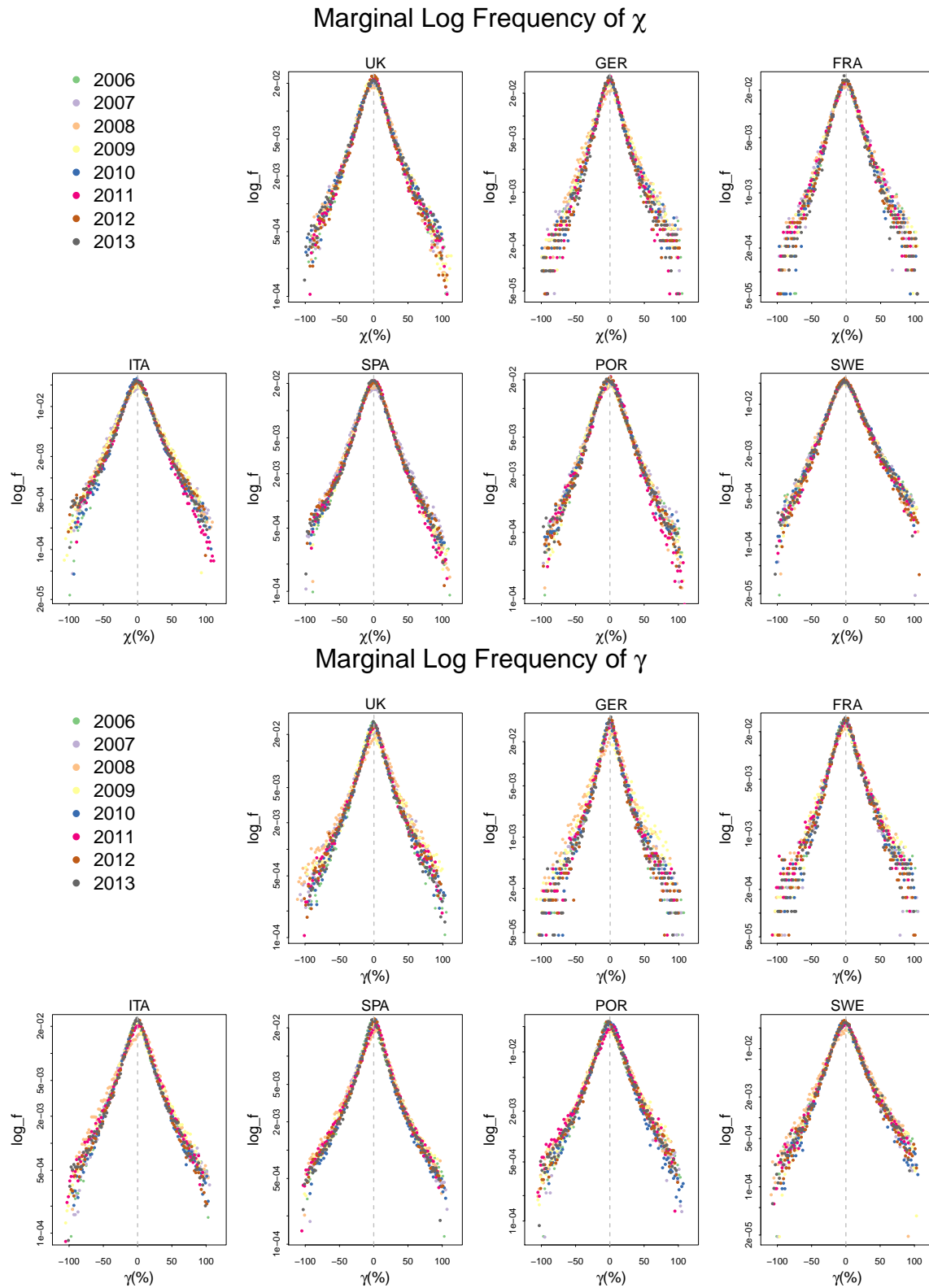
**Figure 2.** The frequency distribution of the growth rate of labor and capital productivity, $\gamma$ and $\chi$, with log scale on the vertical axis. The histograms are centered for each country and each year. Both distributions exhibit a peaked tent shape.

## Marginal Log Frequency of $\zeta$



**Figure 3.** The frequency distribution of the growth rate of cost reduction $\zeta$ with log scale on the vertical axis. The histograms are centered for each country and each year. The distribution exhibits a peaked tent shape.
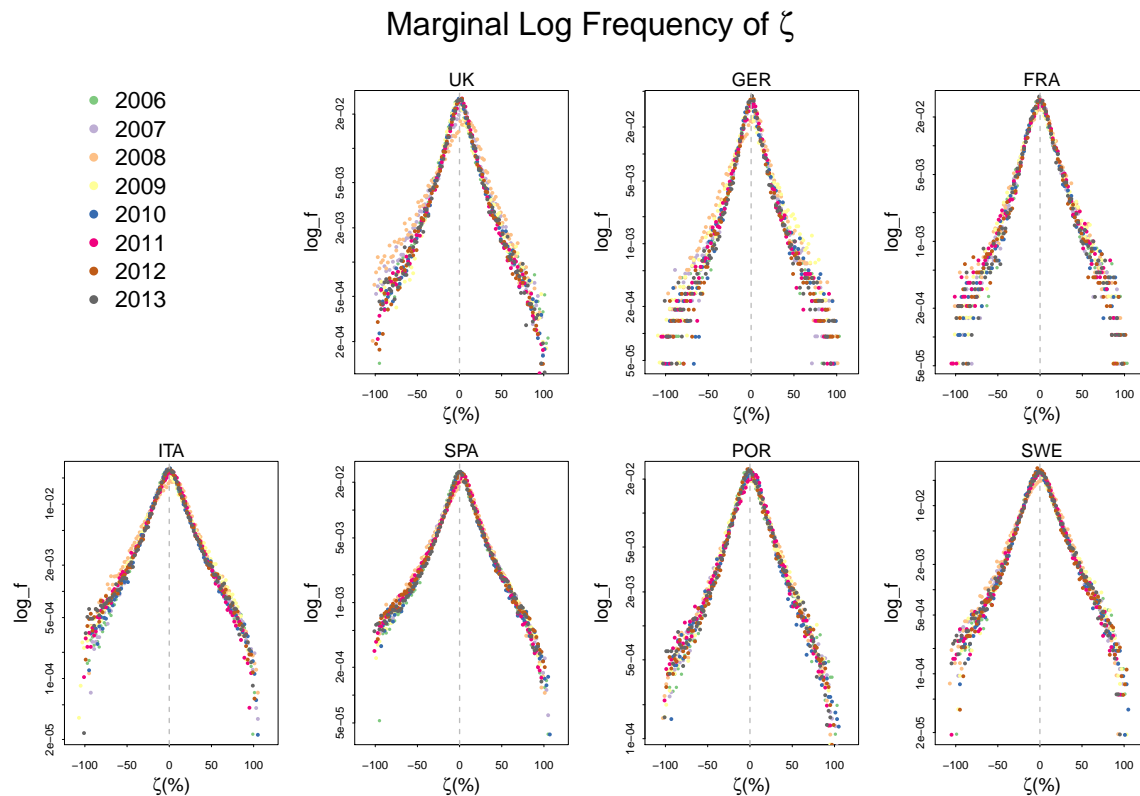
## 3. A Statistical Equilibrium Model of the ITC

A highly peaked and symmetric pattern of the frequency distributions of $\gamma, \chi,$ and $\zeta$ suggest that there is a central tendency in the distribution and structured deviations from it. The observed pattern of technical progress departs from the single-state equilibrium, such as the Walrasian equilibrium, in that the deviation from the central tendency (that is, the mode of the frequency distribution) is persistent and is not properly explained by the normal distribution as is in the case of many stochastic Walrasian models. (The actual fit of the observed data of $\zeta$ is better explained by the Laplace distribution $f(x \mid \mu, b) = \frac{1}{2b} \exp\left(-\frac{|x-\mu|}{b}\right)$, where $\mu$ and $b$ are the location and the scale parameter. There are a few notable economic studies deriving and applying the Laplace distribution from the statistical equilibrium perspective. For example, see [17,18] for their studies on the firm profit rates and growth rates.) A more proper concept of equilibrium compatible with the pattern of the observed data is one that predicts the equilibrium as a distribution of different states.

While the notion of equilibrium as a probability distribution is relatively unfamiliar to economists, it has been widely accepted in physics and information theory under the name of *statistical equilibrium* [1,2]. Statistical equilibrium represents the most likely state of the system in the form of a probability distribution $p(x)$, which can be derived by maximizing the entropy of the system, $H[p(x)] = -\sum p(x) \log[p(x)]$. (For the discussions on the derivation of different statistical equilibrium from the maximum entropy, see [19,20].) The Walrasian single-state equilibrium is an unattainable special case of this equilibrium when only one state is assigned a positive probability. The statistical equilibrium of most systems in reality does not collapse to a degenerate distribution but assigns a positive probability to multiple states. Therefore, introducing statistical equilibrium to economics suggests that the goal of economic models is to find a non-degenerate probability distribution of the target system that explains the central tendency along with inherent fluctuations around it.

Based on the statistical equilibrium approach, we will introduce one class of model called the QRSE model [15]. Suppose a finite set of outcomes $X \rightarrow R$ and a finite discrete set of actions $A = \{a_1, \ldots, a_n\}$. In our example of the firm's choice of technique, the discrete action variable is a binary set $A = \{a, \bar{a}\}$ consisting of two complementary actions with $a$ and $\bar{a}$ indicating the adoption and non-adoption of a new technology, while the set of outcomes consists of the rate of cost reduction expected from the new technology. The key dynamics of this model is that the outcome and action variables interact with each other. First, the quantal response part of the model predicts that the firm decision on the adoption of a new technology is determined in response to the outcome variable, $\zeta$, the degree of potential rate of cost reduction. The impact of the potential cost reduction on the probability of adoption is expressed through the conditional probability of $A$ on $\zeta$, $p(A|\zeta)$. Second, the model predicts that the rate of cost reduction itself is also affected by the firm's act of adopting a new technology. The impact of adoption of technology on outcome variables is expressed as $p(\zeta|A)$.

Different economic theories of technical change can lead to different specifications of these two-way interactions expressed by $p(A|\zeta)$ and $p(\zeta|A)$. The following subsections discuss the QSRE model of ITC to specify these interactions and derive the statistical equilibrium distribution of the rate of cost reduction.

### 3.1. The Impact of the Cost Reduction on Adoption of a New Technology

We employ the induced technical change (ITC) model [21–25] as the baseline behavioral model for a choice of technique. In this model, a typical firm maximizes $\zeta$ by adopting a new technology constrained by the *innovation possibilities frontier* (IPF), which defines a trade-off between increases in labor and capital productivity. To maximize $\zeta$, the firm adopts a technology that affects $\gamma$ and $\chi$ in response to changes in $\omega$ and $\pi$. For example, if there is an increase in unit labor cost $\omega$, the firm will be better off introducing a labor-saving technology, which is expressed by increasing $\gamma$ and decreasing $\chi$.

Following the logic of the ITC model, our model assumes that the probability of the firm adopting a new technique depends on how much cost reduction $\zeta$ it can achieve. Therefore, the quantal response function is expressed by the conditional distribution of the adoption decision on the potential rate of cost reduction, $p[A|\zeta]$. The quantal response function has an associated payoff function $u[A, \zeta]$ representing the payoff to the typical firm of adopting a new technique. The model assumes that the typical firm maximizes the expected payoff with a mixed strategy of $A = \{a, \bar{a}\}$. This boils down to a simple maximization problem as follows:

$$\max \quad \sum_A p(A|\zeta) u(A, \zeta), \tag{1}$$
$$\text{s. t} \quad \sum_A p(A|\zeta) = 1.$$

With no further constraint, the solution to this problem is the Dirac Delta function, choosing to either adopt or not adopt:

$$p(A|\zeta) = \text{DiracDelta}(A - \hat{A}[u, A, \zeta]), \tag{2}$$

where $\hat{A}[u, A, \zeta]$ is the choice of adoption or non-adoption that maximizes the payoff. Therefore, the resulting frequency distribution of $p(A|\zeta)$ puts unit weight on the payoff-maximizing action and zero weight on the other.

From a statistical equilibrium point of view, however, this result is extremely unlikely to happen because it requires the entropy of the system to be zero. In economics, the zero entropy case can be understood as a *perfect rationality model*, in which the typical firm has a full capacity to process all relevant market signals, resulting in a complete certainty about her decision. In the context of technical change, the perfect rationality model implies that any changes in the input costs will induce an optimal response of technical change so that the potential rate of cost reduction is fully exhausted. Since the zero

entropy is not attainable in the real world and remains only as an unrealistic theoretical entity, we need to generalize the model by introducing a positive minimum entropy $H_{min}$. This simple modification is one way to model the *bounded rationality* of the economic agent. Consequently, the model yields a different maximization problem with an entropy constraint as follows:

$$\begin{aligned} \max \quad & \sum p(A|\zeta)u(A|\zeta), \\ \text{s.t} \quad & \sum_a p(A|\zeta) = 1, \\ & H[p(A|\zeta)] = -\sum p(A|\zeta) \operatorname{Log}[p(A|\zeta)] \geq H_{min}. \end{aligned} \tag{3}$$

A Lagrangian function of this maximization problem is:

$$\mathrm{L} = \sum p(A|\zeta)u(A,\zeta) - \mu \left( \sum p(A|\zeta) - 1 \right) + T \left( \sum p(A|\zeta) \operatorname{Log}[p(A|\zeta)] - H_{min} \right).$$

The resulting frequencies of $p(A|\zeta)$ at a behavior temperature or "shadow price" $T$ is:

$$p(A|\zeta) = Z(u,T,A)^{-1} e^{\frac{u(A,x)}{T}} = \frac{e^{\frac{u(A,\zeta)}{T}}}{\sum e^{\frac{u(A,\zeta)}{T}}}, \tag{4}$$

where the partition function $Z(u,T,A) = \sum e^{\frac{u(A,\zeta)}{T}}$. The result suggests that the optimal technical change is not a single rate of cost reduction, but a probability distribution of all the possible rates of cost reduction. The behavior temperature $T$, which was originally called *entropy prices* in Foley's seminal work on the statistical equilibrium approach to economics [4], plays an important economic role because it determines the overall intensity of the payoff of economic actions in the market. As we will see in more detail in the subsequent discussions, higher behavior temperature $T$ implies lower intensity of the payoff and thus more uncertainty over the possible economic actions. This form of Quantal Response (QE) model has been used to model bounded rationality in economics. A general survey of QE models can be found in [26–28]. One major difference of the QRSE model from the previous models in the QE literature is that our QE model is obtained as an implication of the maximum entropy principle, not a behavioral assumption. (For a seminal work on the QE models from the game-theoretic perspectives, see [29–31]. A discussion on the entropy theoretic approach to the bounded rationality can be found in [32]. I appreciate the anonymous referees of this journal for this point.)

Further deriving $p(A|\zeta)$ for a binary variable $A$, we have:

$$\begin{aligned} p(A = a|\zeta) &= \frac{e^{\frac{u(a,\zeta)}{T}}}{e^{\frac{u(a,\zeta)}{T}} + e^{\frac{u(\bar{a},\zeta)}{T}}} = \frac{1}{1 + e^{-\frac{u(a,\zeta)-u(\bar{a},\zeta)}{T}}}, \\ p(A = \bar{a}|\zeta) &= 1 - p(a|\zeta) = \frac{e^{-\frac{u(a,\zeta)-u(\bar{a},\zeta)}{T}}}{1 + e^{-\frac{u(a,\zeta)-u(\bar{a},\zeta)}{T}}}. \end{aligned}$$

The recovered conditional distribution $p(A|\zeta)$ gives the probability of a particular action given observed economic variable $\zeta$. Except for the case when the behavior temperature $T$ is zero, the link function is not degenerate and assigns positive probabilities to heterogeneous responses.

The payoff difference $u(a,\zeta) - u(\bar{a},\zeta)$ (that is, the difference in payoff between adoption and non-adoption) can be modeled to satisfy the following conditions. First, a higher $\zeta$ increases the payoff of adoption while lowering the payoff of non-adoption, and therefore increases the payoff difference. Second, the payoff difference becomes zero so that the firm is indifferent to the adoption of a new technology when $\zeta$ is equal to a shift parameter $\mu$ that determines the indifference point or the hurdle point of the rate of cost reduction where the probability of adopting a new technology is 50%. This shift parameter $\mu$ can be understood as a "premium" on the rate of cost reduction required for the

firm to adopt the technology. If $\mu$ is high, this implies that the firm needs a higher premium on the rate of cost reduction to adopt the technology with a higher than 50% probability. The simplest linear function that reflects these two constraints is:

$$u(a, \zeta) - u(\bar{a}, \zeta) = \zeta - \mu. \tag{5}$$

Therefore, the impact of the rate of cost reduction on adoption of a new technology is modeled as

$$\begin{aligned} p(A = a|\zeta) &= \frac{1}{1 + e^{-\frac{\zeta - \mu}{T}}}, \\ p(A = \bar{a}|\zeta) &= \frac{e^{-\frac{\zeta - \mu}{T}}}{1 + e^{-\frac{\zeta - \mu}{T}}}. \end{aligned} \tag{6}$$

Figure 4 shows the quantal response function with the behavior temperature $T$:
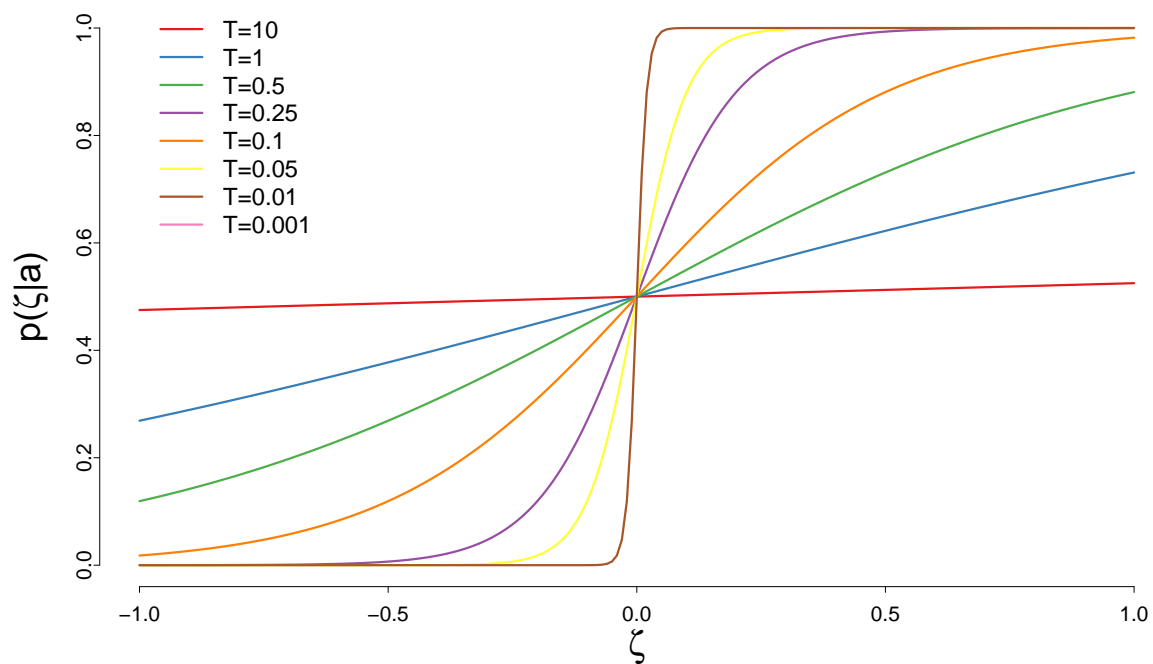
## Logit Quantal Response Function



**Figure 4.** A quantal response function with the behavior temperature $T$ and $\mu = 0$. The horizontal axis represents the the rate of cost reduction, $\zeta$, while the vertical axis represents the frequency of the adoption of a technology.

This figure shows that the higher the behavior temperature, the more uncertain the decision. Except for the unattainable case when $T = 0$, the function predicts a gradual increase in the frequency of action in response to a higher rate of cost reduction. When $T$ is sufficiently large, the quantal response becomes uniform across different actions (the red line). When the behavior temperature $T$ is close to zero, the function becomes a step function (the brown line) [14].

### 3.2. The Impact of the Adoption of a New Technology on the Rate of Cost Reduction

The key assumption of the ITC model is that the firm faces a limit in technical progress, the IPF, through which available technologies are constrained by the trade-off between the rates of increase of labor and capital productivity. Under the IPF constraint, an increase in productivity of one input is made possible at the cost of a decrease in productivity of the other input. For example, higher labor productivity growth is coupled with lower capital productivity growth on the innovation possibilities frontier. The trade-off between $\chi$ and $\gamma$ can be represented by the concave function:

$$\gamma = f(\chi), \quad \text{with} \quad f' < 0, \quad f'' < 0. \tag{7}$$

Figure 5 displays a hypothetical IPF:
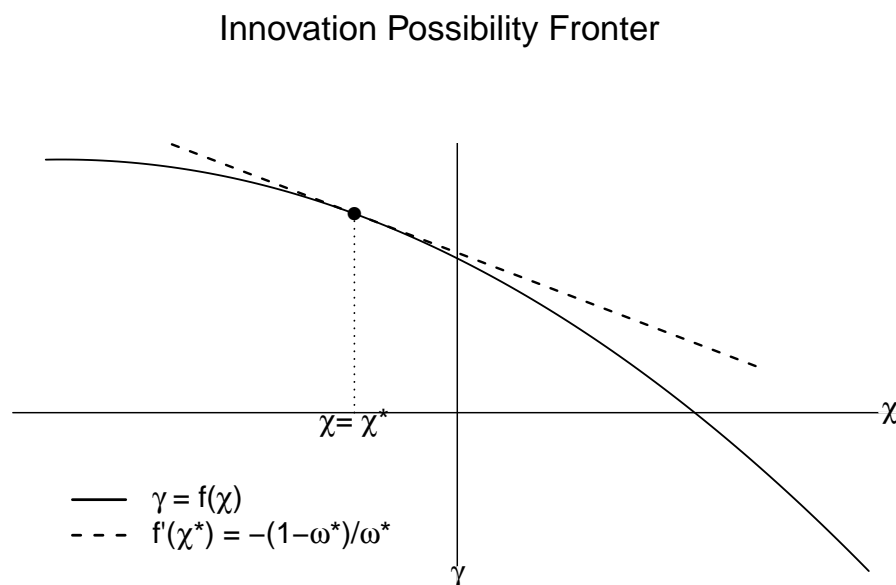
## Innovation Possibility Fronter



**Figure 5.** An innovation possibilities frontier. The bold line represents the innovation possibilities frontier (IPF), a functional relationship of the trade-off between $\chi$ and $\gamma$. The dotted line represents a tangent line of IPF at $\chi = \chi^*$. The tangent is the maximum rate of cost reduction given the unit labor and capital cost $\omega = \omega^*$ and $\pi = \pi^*$.

The key property of the IPF is that the tangent of the function is the negative factor-price ratio, $-\pi/\omega$ [22]. Depending on a particular $\pi/\omega$, the firm chooses the corresponding cost-minimizing techniques on the frontier.

The second primary assumption of the QRSE model of technical change is that the adoption of new technology has an impact on the rate of cost reduction through changes in the factor-price ratio. In the model, the typical firm adopts a new technology with a higher than average rate of cost reduction, and therefore, the expected rate of cost reduction conditional on adoption is greater than the expected rate of cost reduction conditional on non-adoption. Now suppose that the act of adopting a new technology does not make any impact on the factor market. In this case, the difference in the conditional expectations of the rate of cost reduction will not be corrected, so that the firm with a new technology will continue to have a higher rate of cost reduction. However, if the new technology changes the factor prices in the market and makes the input initially saved become more expensive, the initial gain in the increased rate of cost reduction will be lost unless the firm finds another new optimal technology for the changed factor-price ratio.

To illustrate, suppose that the firm adopts a labor-saving and capital-consuming technology in response to a higher cost of labor input relative to the capital input. As a result, the firm would require less labor and more capital, which, in turn, will twist the factor-price ratio against the firm's

initial choice by making the labor input cheaper relative to the capital input. Following [15], we can model this negative feedback (or the competitive pressure) in terms of the difference between the two conditional expectation of $\zeta$ weighted by the marginal probability of action variable:

$$E(\zeta|a)p(a) - E(\zeta|\bar{a})p(\bar{a}) \leq \delta, \tag{8}$$

where $\delta$ represents the degree of the "ineffectiveness" of the competitive pressure. The larger $\delta$ is, the less effective market response is as to the adoption of new technology. When $\delta = 0$, it implies that the factor market is so effective that it completely corrects any impacts from the technical change. In contrast, when $\delta = \infty$, the cost reduction achieved by a new technology will not be offset by the negative feedback from the adverse changes in the factor-price ratio.

*3.3. Maximum Entropy Program of the Quantal Response ITC Model*

Up to now we have discussed two model assumptions in terms of the interaction between the action variable $A$ and the outcome variable $\zeta$. Additionally, we introduce the mean constraint on $\zeta$, $E(\zeta) = \psi$, which represents the typical rate of cost reduction deemed by the cost-minimizing firms in a competitive market. (For a detailed discussion on the first moment constraints in economic models, see [33].) These assumptions can be used as the constraints of our maximum entropy program to find the statistical equilibrium of $A$ and $\zeta$. Using the constraints in Equation (7) and (8), and the mean constraint of the model $E(\zeta) = \psi$, the maximum entropy program can be written as follows:

$$\max \quad -\int \sum_a p(A,\zeta) \operatorname{Log}[p(A,\zeta)]d\zeta, \tag{9}$$

$$\text{s. t} \quad \int \sum_a p(A,\zeta)d\zeta = 1$$

$$\int p(\zeta)\zeta d\zeta = \psi$$

$$\int p(a,\zeta)\zeta d\zeta - \int p(\bar{a},\zeta)\zeta d\zeta \leq \delta$$

$$p(A = a|\zeta) = \frac{1}{1 + e^{-\frac{\zeta-\mu}{T}}}.$$

The solution to this problem is expressed in terms of the joint distribution $\hat{p}(A,\zeta)$ which implies two "predicted" marginal distributions, $\hat{p}(A), \hat{p}(\zeta)$, and two conditional distributions, $\hat{p}(A|\zeta), \hat{p}(\zeta|A)$. Since the data on the action variable $A$ are not observable, we need to express the solution in terms of the marginal distribution of $\zeta$, $\hat{p}(\zeta)$ to be able to estimate the unknown parameters of the model. The solution to this program then becomes:

$$\hat{p}(\zeta) = \frac{e^{H_{\mu,T}} e^{-\beta \operatorname{Tanh}\left[\frac{\zeta-\mu}{2T}\right]\zeta} e^{-\kappa\zeta}}{\sum_\zeta e^{H_{\mu,T}} e^{-\beta \operatorname{Tanh}\left[\frac{\zeta-\mu}{2T}\right]\zeta} e^{-\kappa\zeta}}, \tag{10}$$

where $H_{\mu,T}(\zeta) = H\left[\frac{1}{e^{-\frac{\zeta-\mu}{T}}+1}, \frac{1}{e^{\frac{\zeta-\mu}{T}}+1}\right]$ and $\operatorname{Tanh}[\alpha]$ is the hyperbolic tangent function, written as $\frac{e^{2\alpha}-1}{e^{2\alpha}+1}$. (See [15] for the proof. The solution is in discrete terms and requires the coarse-grained bins of $\zeta$.) There are four unknown parameters, $\mu, T, \beta$ and $\kappa$. $\mu$ and $T$ are the location and the shape parameter of the quantal response function, $\beta$ is the Lagrangian multiplier of the $\delta$ constraint and implies the impact of adoption $A$ on the cost reduction $\zeta$, and $\kappa$ is the Lagrangian multiplier of the mean constraint $\psi$, which determines the skewness of the distribution. A graphical characterization of distribution (10) is provided in Figure 6.
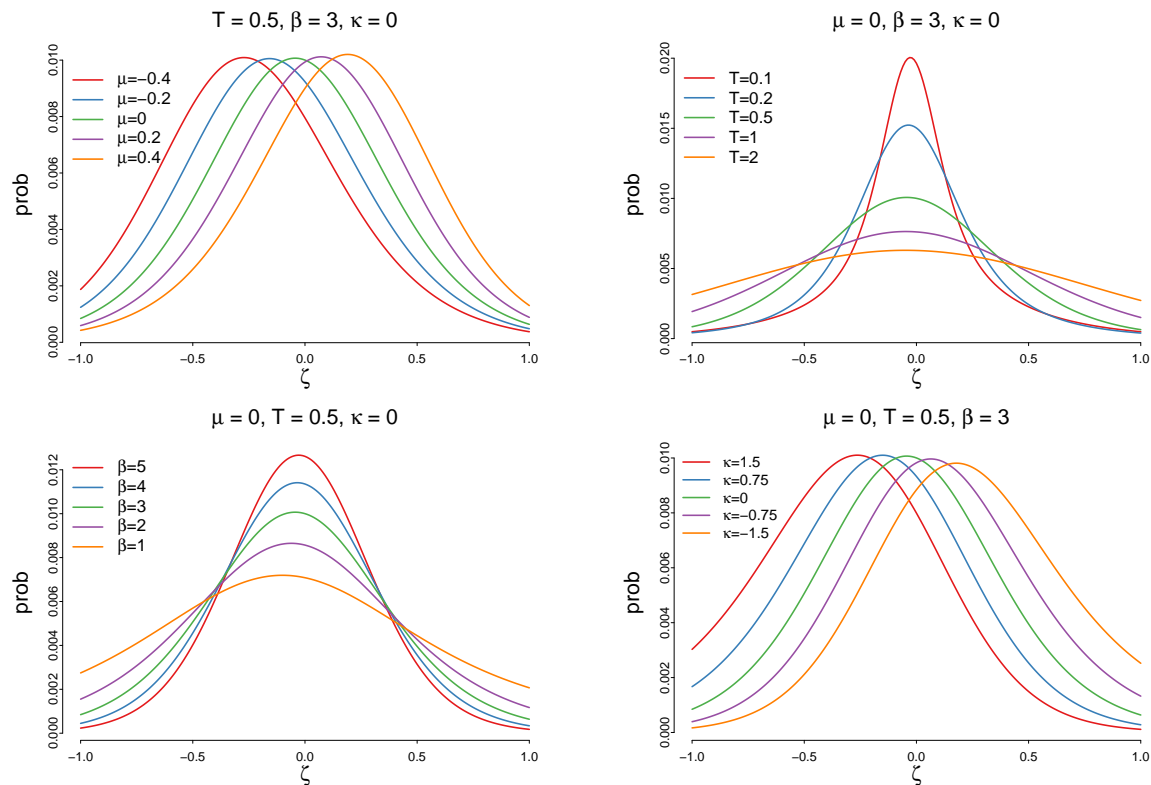
## QRSE Distributions



**Figure 6.** QRSE distributions with different parameter values of $\mu, T, \beta$, and $\kappa$. $\mu$ determines the location of the distribution, while $\kappa$ determines its skewness. $\kappa$ close to zero represents a symmetric QRSE distribution. Predictably, $T$ and $\beta$ determine how disperse the distribution is. The lower $T$ and higher $\beta$ are, the more spread out the distribution becomes.

## 4. Bayesian Estimation of the Model

### 4.1. Model Specification

In estimating $\mu, T, \beta$ and $\kappa$, we will rely on the Bayesian inference. In setting up the likelihood function, we utilize the observed distribution of $\zeta$ and use it as a reference distribution whose divergence from $\hat{p}(\zeta)$ determines the likelihood of the four unknown parameters. (One could use the derived maximum entropy distribution as a likelihood function and evaluate it directly. This method requires a numerical integration of the partition function since it does not have a closed form solution. One small issue with this method is that the integration needs to be repeated at each evaluation, which can slow down the computation, especially when we sample from it with a large number of iterations and multiple chains in the simulation.) For the comparison of the predicted and the observed distribution, we use the Kullback-Leibler divergence [34]. The KL divergence measures a discrepancy between two distributions and thus has been widely used as a measure of comparing different distributions in the context of many statistical inferences such as model estimation, model selection, and classification. The KL divergence of the discrete distribution of $p$ given a known distribution $q$, $D_{\text{KL}}(p\|q)$, is defined as follows:

$$
\begin{aligned}
D_{\text{KL}}(p\|q) &= \sum_i p(x_i) \log \frac{p(x_i)}{q(x_i)} \\
&= H(p,q) - H(p).
\end{aligned} \tag{11}
$$

Regarding the goodness of fit test, the KL divergence measures how close the distribution of actual data $p(x)$ is to the candidate model of probability distribution $q(x)$. A lower KL divergence implies that it is more likely that the observed distribution of $p(x)$ comes from the presumed model $q(x)$.

Using the KL as a metric, we set up a likelihood function for a lower KL divergence to have a higher likelihood in the function. Any monotonically decreasing function of $D_{KL}$ can be a candidate. However, the KL divergence itself is actually a good approximation to the log-likelihood for the multinomial model as follows:

$$\log[p(X|p)] \approx -n * D_{KL}(\mathbf{X}/n||\mathbf{p}),\tag{12}$$

where $\mathbf{X} = (X_1, X_2, ..., X_k)$ is the random variable that represents the number of times the outcome $k$ occurs, $\mathbf{p} = (p_1, p_2, ..., p_k)$ is the probability of outcome $k$, and $n$ is the total number of observations. Since we have the coarse-grained bins of observed distribution of $\zeta$, which can be interpreted as a particular realization of multinomial distribution, we can think of the frequency of each outcome as a sample of a multinomial model with the predicted probability of $\zeta$ from Equation (10). Therefore, denoting the coarse-grained bin of $\zeta$ as $\bar{p}(\zeta)$ and remembering that $\hat{p}$ is the maximum entropy distribution, the log-likelihood function becomes:

$$\log[p(\mu, T, \beta, \kappa|\hat{p})] = -n * D_{KL}(\bar{p}(\zeta)|\hat{p}(\zeta)).\tag{13}$$

### 4.2. Result

We evaluate $\log[p(\mu, T, \beta, \kappa|\hat{p})]$ in Equation (14) using a sequence of random samples from the posterior distribution obtained from the Metropolis-Hastings (MH) algorithm, one of Markov Chain Monte Carlo (MCMC) methods. We use 50,000 iterations and 3 chains to recover the marginal posterior distributions of parameters $\mu, T, \beta, \kappa$ for 7 countries in 8 different years. For all the parameters $\mu, T, \beta$ and $\kappa$, we use uninformative priors (uniform priors). For detailed discussions on the MCMC methods, see [35–37]. (The result is not greatly sensitive to a wide range of different binning schemes since we have a large number of data points within a relatively small range $-100\%$ and $-100\%$.)

Parameter Estimation

Figure 7 summarizes the recovered coefficients after 25,000 burn-in periods of the simulation along with 95% credible interval. (All three chains in the simulations are well mixed and properly converge with $\hat{R}$ being 1 in all cases. For a detailed discussion on the convergence diagnosis using $\hat{R}$, see [37].)
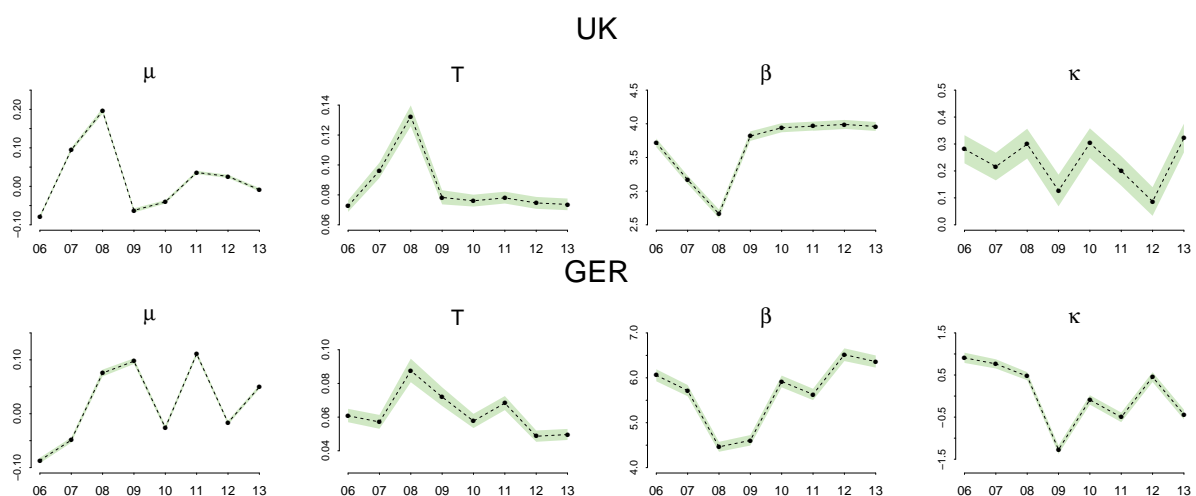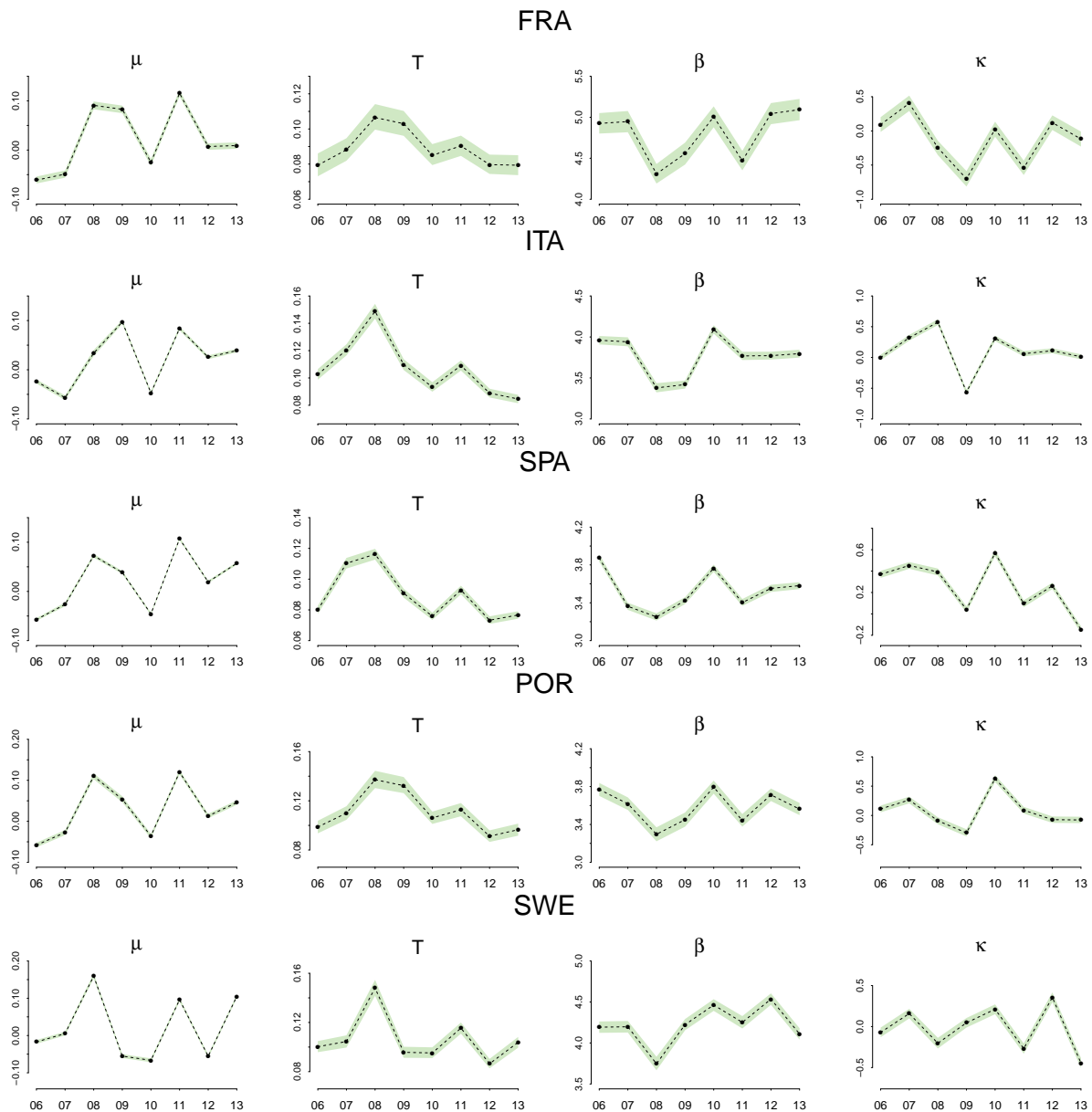


**Figure 7.** *Cont.*

**Figure 7.** Estimated parameters $\mu, T, \beta,$ and $\kappa$ along with the 95% credible interval. The year index represents the beginning year from which the growth rate of $\zeta$ is calculated. For example, the growth rate in year 2006 is the growth of $\zeta$ between 2006–2007.

The 95% credible interval of the parameters show that the posterior distribution has a very small standard deviation due to the large sample size. An increase in $\mu$ and $T$ and a decrease in $\beta$ are noticeable during the 2007–2009 financial crisis. $\kappa$ is close to zero with no clear patterns. The interpretation of this result is discussed in detail in the following sections.

Comparison of Prior and Posterior Distribution: UK 2011

We can visualize the posterior distributions of the parameters using the estimation from the UK in 2011 as a representative case whose result is more or less the same as other estimations regarding the shape of posterior distribution in Figure 8.
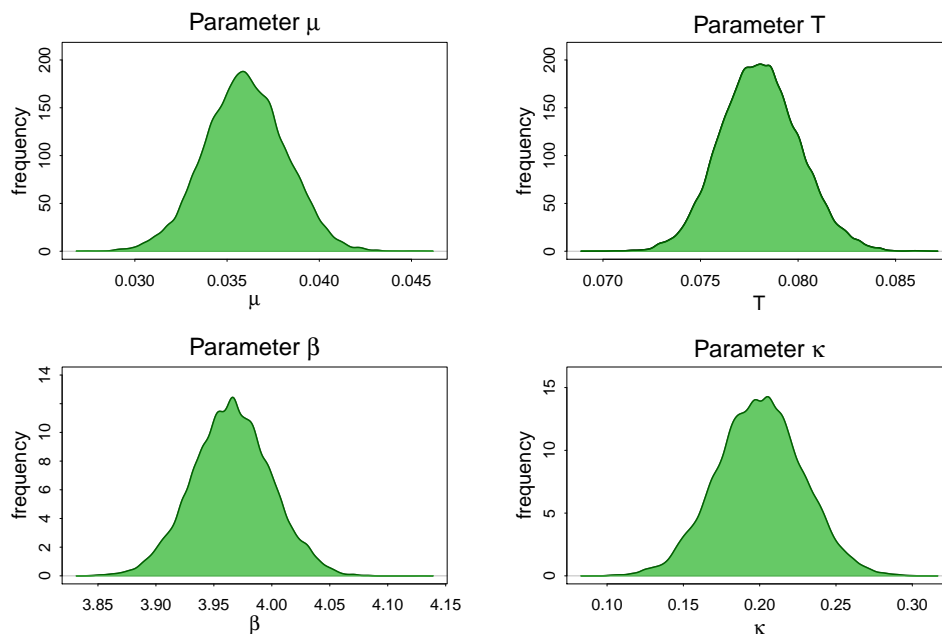
## Prior/Posterior Distributions: UK_2011



**Figure 8.** The posterior distributions of $\mu, T, \beta$ and $\kappa$.

The posterior distributions of all parameters $\mu, T, \beta$ and $\kappa$ are well recovered and exhibits a unimodal symmetric distribution with a very small standard deviation.

Predicted $\tilde{p}(\zeta)$, $\tilde{p}(A|\zeta)$, and $\tilde{p}(\zeta|A)$: UK 2011

Using the estimated parameters, we can recover the marginal and the conditional distributions: $\tilde{p}(\zeta)$, $\tilde{p}(A|\zeta)$, and $p(\tilde{\zeta}|A)$. Figure 9 visualizes the recovered distribution, again using the estimation of the UK in 2011 as a representative example.

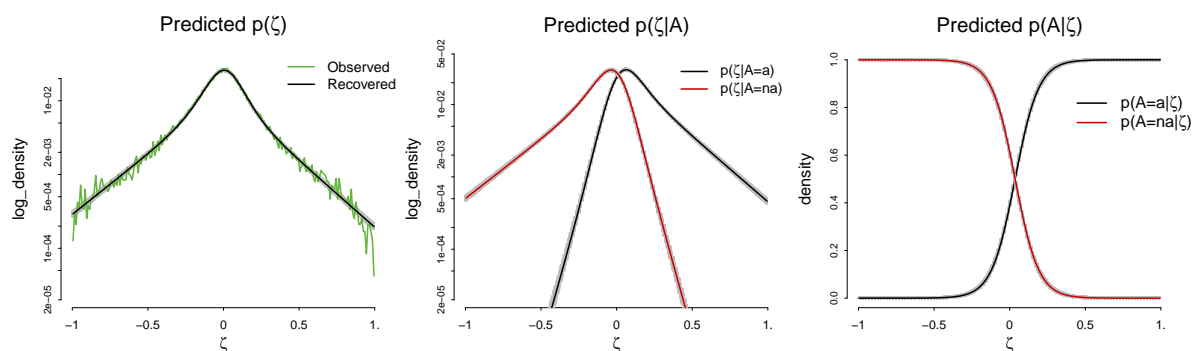## Recovered Distributions: UK_2011



**Figure 9.** The recovered distribution of $\tilde{p}(\zeta)$, $\tilde{p}(A|\zeta)$, and $\tilde{p}(\zeta|A)$. Distributions are recovered using the mean value of estimated parameters along with 95% credible interval expressed as the scattered gray points.

The first panel in Figure 9 represents the recovered marginal distribution of $\zeta$, $\tilde{p}(\zeta)$, along with the observed one, $\bar{p}(\zeta)$. The black line is $\tilde{p}(\zeta)$ recovered from the mean value of estimated parameters, while the green line is $\bar{p}(\zeta)$. The visual inspection shows that the model satisfactorily recovers the observed distribution. The informational performance of the fitted distribution based on the Soofi ID

shows that $\tilde{p}(\zeta)$ explains 99.6% of the informational content of $\bar{p}(\zeta)$. (A detailed discussion on the Soofi ID and the informational performance of all the fitted distributions can be found in Appendix.) The second panel shows the recovered conditional distribution of $\zeta$ on the adoption decisions, $\tilde{p}(\zeta|A)$. $\tilde{p}(\zeta|A = a)$ is skewed to the right while $\tilde{p}(\zeta|A = \bar{a})$ is skewed to the left, implying that the former has a higher expected value of $\zeta$. This confirms the initial model assumption that the conditional expectation of $\zeta$ on adoption, $E(\zeta|A = a)$, is greater than conditional expectation of $\zeta$ on non-adoption, $E(\zeta|A = \bar{a})$ constrained by $\delta$. The final panel displays the recovered quantal response function, the conditional distribution of the adoption decisions on $\zeta$, $\tilde{p}(A|\zeta)$. A substantial degree of uncertainty exists in adopting a new technology as is shown by the fact that the recovered link function is far from the step function.

### 4.3. Discussion of the Estimated Parameters, $\mu, T, \beta$ and $\kappa$

The posterior estimation of $\mu, T, \beta$ and $\kappa$ for seven countries over eight years summarized in Figure 7 exhibits four important patterns.

First, the estimated $\mu$ sharply increases during the financial crisis 2007–2009 in all countries. As we discussed before, $\mu$ represents the location parameter of the quantal response function that determines the indifference point between adoption and non-adoption. A positive $\mu$ means that a positive premium on the potential rate of cost reduction is required for the firm to adopt a new technology with a higher than 50% probability. A sharp increase in $\mu$ during the financial crisis demonstrates that the firms became more reluctant to adopt a new technology and needed a higher premium to do so. This point can be more clearly demonstrated by taking the difference between the recovered $\mu$ and the observed mean of $\hat{\zeta}$ summarized as follows:

Figure 10 shows that, before the financial crisis in 2006, $\mu$ is almost equal to (or slightly smaller than) the average $\zeta$ in all countries, implying that the firms did not need an additional premium on the rate of cost reduction above the social average to adopt a new technology. During the financial crisis, however, $\mu - \hat{\zeta}$ sharply increases and turns positive, requiring higher than the average premium for the firms to adopt a new technology with a higher than 50% probability.
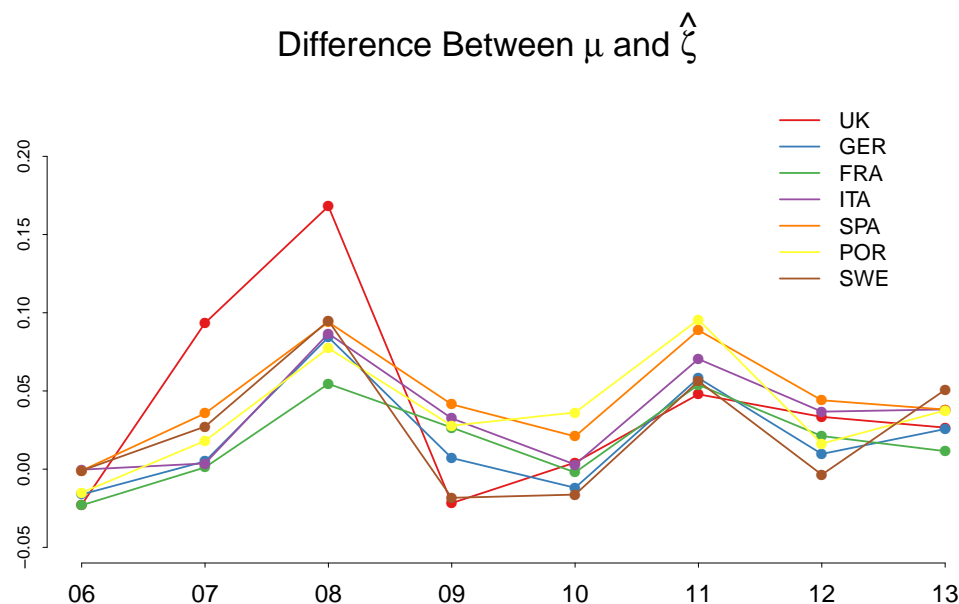


**Figure 10.** The difference between the recovered $\mu$ and the observed mean of $\hat{\zeta}$, $\mu - \hat{\zeta}$.

Second, the estimated $T$ also increases during the crisis period. The parameter $T$ is the shape parameter of the quantal response function that determines the spread of the function. The higher $T$ is, the more spread out the response function becomes, making the probability of adoption and

non-adoption less distinguishable. In economic terms, the spread of the function is interpreted as a degree of the bounded rationality of the economic agent and therefore quantifies how uncertain the economic agent is about her decision. In this regard, an increase in $T$ in 2007–2009 can be explained in terms of a higher degree of uncertainty in adopting a new technology during the crisis. This implies that the firms more (or less) frequently adopted those technologies that adversely (or favorably) affected the cost efficiency, incurring more economic losses than other periods.

Third, $\beta$ decreases during the crisis period. In the maximum entropy program, $\beta$ is the Lagrangian multiplier of the $\delta$ constraint that limits the impact of adoption on the rate of cost reduction. The larger the $\beta$ is, the smaller the $\delta$, and therefore, the more effective the impact of adoption on the potential rate of cost reductions. The fact that $\beta$ was lower during the financial crisis suggests that the market effectiveness was lower during the market turmoil than other periods, and therefore, the technological progress of the firms didn't effectively impact the factor-price on the market.

Fourth, the skewness parameter $\kappa$ is close to zero and has no clear patterns in most countries, except for Germany and France which have seen a drop in $\kappa$ during the crisis. In the maximum entropy program, $\kappa$ is the Lagrangian multiplier of the mean constraint $\psi$. The fact that $\kappa$ is close to zero means that the distributions of $\zeta$ have been relatively symmetric on the whole.

## 5. Discussion

From the statistical inference perspective, the initial set-up of our model in terms of the joint distribution of $\zeta$ and $A$ is ill-posed due to the fact that the variable $A$ is not observable. The MEP can provide a systematic solution to this problem of limited data [38,39]. The advantage of the MEP method comes from its core inferential logic that the maximum entropy distribution implies the most likely state of the system, regardless of data availability. With adequate economic theories as constraints, the MEP derives the "prior distribution of data", with some unknown parameters. This predicted distribution can be fitted to the observed data in the way that minimizes the model's informational loss (the KL divergence). Recovered unknown parameters can be used to infer the unobservable variables, that is, in our case, the adoption decision $A$. This enables us to recover latent distributions involving unobservable variable, such as $p(A|\zeta)$ and $p(\zeta|A)$, providing a rich picture of the system of economic dynamics.

Another important aspect of our statistical model of technical progress is that it is fundamentally different from the conventional Walrasian equilibrium models. Both models assume that the social outcome is a result of a micro-behavior of the economic agent. However, unlike the "representative agent" in the Walrasian equilibrium model, our model does not assume that the agents are identical. The "typical agent" in our model has a quantal response function with a non-zero behavior temperature as a behavioral rule. Even with the same payoff function, the quantal response makes possible a wide range of heterogenous behaviors. (The statistical model of technical progress is not a single behavioral model but a whole class of different behavioral models. This is because it inherently abstracts from any behavioral assumptions but still predicts the heterogeneity of the economic results due to the non-zero behavior temperature $T$ of the quantal response function.) Also, in the Walrasian equilibrium model, the economic agent is inherently atomistic in the sense that she does not make any impact on the aggregate economic outcome. Our model relaxes this assumption and allows a positive degree of interaction so that the typical agent exerts an impact on the aggregate outcome.

## 6. Conclusions

This paper has developed a QRSE model of induced technical change with the binary action variable (adoption or non-adoption of a new technology) and the outcome variable (the rate of cost reduction). Constraints are put on the interactions of the action- and outcome-based ITC model. The impact of cost reduction on the adoption is modeled using the payoff function of the firm, which assumes that higher rates of cost reduction yield higher payoffs with a positive degree of uncertainty in the decision process. The impact of the adoption of new technology on the rate of cost reduction

is formulated utilizing the idea of the economic impact of input-saving technical progress on the factor-price ratio.

The principle of maximum entropy is then applied to find the statistical equilibrium of $\zeta$ and $A$ given four parameters, $\mu, T, \beta$ and $\kappa$ of the model. The recovered joint distribution of $\zeta$ and $A$, which implies the marginal distribution of $\zeta$, is compared to the observed distribution of $\zeta$ using the relative entropy, to set up a likelihood function for unknown parameters. Using the uniform prior distribution, the posterior distributions of the parameters are estimated using the MCMC simulation whose 95% credible intervals are used to recover relevant distributions of our interest, that is, the marginal distribution of $\zeta$, $p(\zeta)$, and the conditional distributions, $p(A|\zeta)$ and $p(\zeta|A)$.

Recovered parameters and distributions imply some interesting features in the patterns of firms' technical progress in the seven advanced OECD countries. During the financial crisis in 2007–2009, the firms became more reluctant to adopt a new technology and needed a higher premium on the potential rate of cost reduction to do so, had a higher degree of uncertainty in adopting a new technology, and the technological progress didn't effectively impact the factor-price on the market.

**Conflicts of Interest:** To the best of my knowledge, I have no conflicts of interest to disclose. This article does not contain any studies with human participants or animals performed by any of the authors.

## Appendix A. Summary Statistics of the Frequency Distributions of $\gamma$, $\chi$, and $\zeta$

The following Table A1 summarizes the mean of $\gamma, \chi, \hat{k}$, and $\zeta$. $\hat{k}$ is defined as $\gamma - \chi$ and represents the change in capital intensity $L/K$.

**Table A1.** The mean of $\gamma, \chi, \hat{k}$ and $\zeta$ of 7 different countries averaged from 2006–2013.

| Country | UK | GER | FRA | ITA | SPA | POR | SWE |
|---|---|---|---|---|---|---|---|
| $\gamma$ (%) | 0.98 | 0.43 | 1.15 | −0.80 | −1.02 | 0.65 | 1.10 |
| $\chi$ (%) | −6.22 | −1.13 | −1.89 | −3.38 | −6.57 | −4.43 | −2.12 |
| $\hat{k}$ (%) | 7.20 | 1.56 | 3.04 | 2.58 | 5.55 | 5.08 | 3.22 |
| $\zeta$ (%) | −2.10 | −0.08 | 0.33 | −1.49 | −2.46 | −0.91 | −0.24 |

The summary statistics show that there has been a labor-saving technical change between 2006 and 2013 in all but Italy and Spain, as is demonstrated by the positive $\gamma$ and the negative $\chi$. The capital intensity has dramatically increased during this period. Except for France, $\zeta$ is shown to be negative implying that firms in those countries have, on average, failed to reduce costs. This unfavorable situation is caused by relatively low (or even negative) growth rate of labor productivities along with a fairly large decrease in the growth rate of capital productivity, especially during the financial crisis in 2007–2009 that adversely affected the firm's revenue.

## Appendix B. Soofi ID Index

Soofi et al. [40] and Soofi and Retzer [41] develop information distinguishability (ID), in which the actual distribution $f(x)$ is compared against the maximum entropy distribution of $x$ as the candidate modle $f^*(x)$. The ID index is defined by:

$$ID(f : f^*|\theta) = 1 - \exp[-D_{\mathrm{KL}}(f\|f^*|\theta)], \tag{A1}$$

where $f$ is the actual data distribution and $f^*$ is the maximum entropy distribution of $x$ with the corresponding parameters $\theta$. The ID index is a normalized measure so $0 \leq ID(f\|f^*|\theta) \leq 1$.

As the KL divergence, a smaller value of ID index implies that the actual distribution is similar to the maximum entropy distribution. When $ID_{(f\|f^*|\theta)} \approx 0$, $f$ and $f^*$ are approximately same Soofi et al. [40]. Finally, different distributions $f_1(x|\theta)$, $f_2(x|\theta)$,..., $f_n(x|\theta)$ are ID distinguishable if $ID_{(f_1\|f^*|\theta)} \neq ID_{(f_2\|f^*|\theta)} \neq, ..., \neq ID_{(f_n\|f^*|\theta)}$ (Soofi, 1995). The following Table A2 summarizes the Soofi ID index of all the fitted distributions of $\tilde{p}(\zeta)$

**Table A2.** The Soofi ID index of all the fitted distributions of $\tilde{p}[\zeta]$ for all 7 countries over 8 years.

| Year | UK | GER | FRA | ITA | SPA | POR | SWE |
|------|------|------|------|------|------|------|------|
| 2006 | 0.006 | 0.069 | 0.043 | 0.006 | 0.010 | 0.010 | 0.006 |
| 2007 | 0.020 | 0.056 | 0.060 | 0.008 | 0.010 | 0.007 | 0.011 |
| 2008 | 0.061 | 0.061 | 0.029 | 0.025 | 0.010 | 0.009 | 0.015 |
| 2009 | 0.007 | 0.050 | 0.039 | 0.020 | 0.003 | 0.012 | 0.008 |
| 2010 | 0.007 | 0.027 | 0.035 | 0.008 | 0.013 | 0.018 | 0.015 |
| 2011 | 0.004 | 0.014 | 0.019 | 0.004 | 0.003 | 0.007 | 0.007 |
| 2012 | 0.008 | 0.017 | 0.018 | 0.003 | 0.002 | 0.009 | 0.004 |
| 2013 | 0.005 | 0.015 | 0.027 | 0.004 | 0.005 | 0.008 | 0.008 |

## References

1. Jaynes, E.T. Where Do We Stand on Maximum Entropy? In *The Maximum Entropy Formalism*; Levine, R.D., Tribus, M., Eds.; MIT Press: Cambridge, MA, USA, 1978; pp. 15–118.
2. Jaynes, E.T. Information theory and statistical mechanics. *Phys. Rev.* **1957**, *106*, 620–630.
3. Farjoun, E.; Machover, M. *Laws of Chaos: Probabilistic Approach to Political Economy*; Verso Books: New York, NY, USA, 1983.
4. Foley, D.K. A Statistical Equilibrium Theory of Market. *J. Econ. Theory* **1994**, *62*, 321–345.
5. Stutzer, M.J. A Bayesian Approach to Diagnostic of Asset Pricing Models. *J. Econ.* **1995**, *68*, 369–397.
6. Stutzer, M.J. Simple Entropic Derivation of a Generalized Black-Scholes Option Pricing Model. *Entropy* **2010**, *2*, 70–77.
7. Sims, C.A. Implications of rational inattention. *J. Monetary Econ.* **2003**, *50*, 665–690.
8. Smith, E.; Foley, D.K. Classical thermodynamics and economic general equilibrium theory. *J. Econ. Dyn. Control* **2008**, *32*, 7–65.
9. Toda, A.A. Existence of a statistical equilibrium for an economy with endogenous offer sets. *Econ. Theory* **2010**, *45*, 379–415.
10. Toda, A.A. Bayesian general equilibrium. *Econ. Theory* **2015**, *58*, 375–411.
11. Lux, T. Applications of statistical physics in finance and economics. In *Handbook on Complexity Research*; Rosser, J.B., Ed.; Edward Elgar: Cheltenham, UK, 2009; pp. 213–258.
12. Zhou, R.; Cai, R.; Tong, G. Applications of Entropy in Finance: A Review. *Entropy* **2013**, *15*, 4909–4931.
13. Rosser, J.B. Entropy and econophysics. *Eur. Phys. J. Spec. Top.* **2016**, *225*, 3091–3104.
14. Yang, J. Information theoretic approaches in economics. *J. Econ. Surv.* **2017**, doi:10.1111/joes.12226.
15. Scharfenaker, E.; Foley, D.K. Quantal Response Statistical Equilibrium in Economic Interactions: Theory and Estimation. *Entropy* **2017**, *19*, 444.
16. Michl, T.; Foley, D.K. *Growth and Distribution*; Harvard University Press: London, UK, 1999.
17. Alfarano, S.; Milaković, M. Does Classical Competition Explain the Statistical Features of Firm Growth? *Econ. Lett.* **2008**, *101*, 272–274.
18. Alfarano, S.; Milaković, M.; Irle A.; Kauschke. J. A Statistical Equilibrium Model of Competitive Firms. *J. Econ. Dyn. Control* **2012**, *36*, 136–149.
19. Jaynes, E.T. *Probability Theory: The Logic of Science*; Cambridge University Press: Cambridge, UK, 2003.
20. Kapur, J.N. *Maximum-Entropy Models in Science and Engineering*; John Wiley & Sons: Hoboken, NJ, USA, 1998.
21. Von Weisäcker, C.C. A New Technical Progress Function (1962). *Ger. Econ. Rev.* **2010**, *11*, 248–265.
22. Kennedy, C. Induced Bias in Innovation and the Theory of Distribution. *Econ. J.* **1964**, *74*, 541–547.
23. Samuelson, P.A. A Theory of Induced Innovation along Kennedy-Weisäcker Lines. *Rev. Econ. Stat.* **1965**, *47*, 343–356.

24.  Drandakis, E.M.; Phelps, E.S. A Model of Induced Invention, Growth and Distribution. *Econ. J.* **1966**, *76*, 823–840.

25.  Foley, D. Endogenous Technical Change with Externalities in a Classical Growth model. *J. Econ. Behav. Organ.* **2003**, *52*, 167–189.

26.  McFadden, D.L. Quantal Choice Analaysis: A Survey. *Ann. Econ. Soc. Meas.* **1976**, *5*, 363–390.

27.  McFadden, D.L. Economic choices. *Am. Econ. Rev.* **1976**, *91*, 351–378.

28.  Train, K. *Discrete Choice Methods with Simulation*; Cambridge University Press: Cambridge, UK, 2009.

29.  McKelvey, R.D.; Palfrey, T.R. Quantal Response Equilibria for Normal Form Games. *Games Econ. Behav.* **1995**, *10*, 6–38.

30.  Chen, H.C.; Friedman, J.W.; Thisse, J.F. Boundedly Rational Nash Equilibrium: A Probabilistic Choice Approach. *Games Econ. Behav.* **1997**, *18*, 32–54.

31.  McKelvey, R.D.; Palfrey, T.R. Quantal Response Equilibria for Extensive Form Games. *Exp. Econ.* **1998**, *1*, 9–41.

32.  Wolpert, D. Information Theory—The Bridge Connecting Bounded Rational Game Theory and Statistical Physics. In *Complex Engineered Systems*; Braha D., Minai A., Bar-Yam, Y., Eds.; Springer: Berlin, Heidelberg, 2006; pp. 262–290.

33.  Dos Santos, P.L. The Principle of Social Scaling. *Complexity* **2017**, *2017*, 8358909.

34.  Kullback, S.; Leibler, R. On information and sufficiency. *Ann. Math. Stat.* **1951**, *22*, 79–86.

35.  Gamerman, D.; Lopes, H.F. *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*, 2nd ed.; CRC Press: Boca Raton, FL, USA, 2008.

36.  Minh, D.D.L.; Minh, D.L. Understanding the Hastings Algorithm. *Commun. Stat. Simul. Comput.* **2015**, *44*, 332–349.

37.  Gelman, A.; Carlin, J.B.; Stern, H.S.; Dunson, D.B.; Vehtari, A.; Rubin, D.B. *Bayesian Data Analysis*, 3rd ed.; CRC Press: Boca Raton, FL, USA, 2013.

38.  Golan, A. Information and Entropy Econometrics—Editor's View. *J. Econ.* **2002**, *107*, 1–15.

39.  Golan, A.; Judge, G.G.; Miller, D. *Maximum Entropy Econometrics: Robust Estimation with Limited Data*; Wiley: Hoboken, NJ, USA, 1996.

40.  Soofi, E.S.; Ebrahimi, N.; Habibullah, M. Information Distinguishability with Application to Analysis of Failure Data. *J. Am. Stat. Assoc.* **1995**, *90*, 657–668.

41.  Soofi, E.; Retzer, J. Information indices: Unification and applications. *J. Econ.* **2002**, *107*, 17–40.