REGULAR ARTICLE

Stochastic macro-equilibrium: a microfoundation for the Keynesian economics

Hiroshi Yoshikawa

Abstract In place of the standard search equilibrium, this paper presents an alternative concept of stochastic macro-equilibrium based on the principle of statistical physics. This concept of equilibrium is motivated by unspecifiable differences of economic agents and the presence of all kinds of micro shocks facing them. Our model mimics the empirically observed distribution of labor productivity. The distribution of productivity resulting from the matching of workers and firms depends crucially on aggregate demand. When aggregate demand rises, not only the unemployment rate declines, but more workers are employed by firms with higher productivity. The effect of the reservation wage on unemployment also depends on aggregate demand so that the distinction between cyclical and structural unemployment is ambiguous. The model, a general equilibrium model of monopolistic competition with friction and uncertainty provides a micro-foundation for Keynes' principle of effective demand.

Keywords Heterogeneous agents · Microeconomic foundations · Unemployment · Productivity · Entropy · Keynes' Principle of effective demand

JEL Classification E10 · E12 · J60

H. Yoshikawa (✉)
Faculty of Economics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-0033, Japan
e-mail: yoshikawa@e.u-tokyo.ac.jp

1 Introduction

The purpose of this paper is to present a new concept of stochastic macro-equilibrium which provides a micro-foundation for the Keynesian theory of effective demand. Cyclical changes of aggregate economic activity, namely quarter to quarter or year to year changes of real GDP are basically determined by changes of aggregate demand. This is the central message of Keynes (1936). Keynes argued that *real* demand rather than factor endowment and technology determines the level of aggregate production in the short-run because the rate of utilization of production factors such as labor and capital endogenously changes responding to changes in real demand. Keynes maintained that this proposition holds true regardless of flexibility of prices and wages; he, in fact, argued that a fall of prices or wages would aggravate, not alleviate the problems facing the economy in deep recession because it may lower aggregate demand. Following Tobin (1993), let us call this proposition the Old Keynesian view.

The challenge is to clarify the market mechanism by which aggregate demand conditions the allocation of production factors in such a way that total output follows changes in real aggregate demand. A decrease of aggregate output is necessarily accompanied by lower utilization of production factors, and *vice versa*. Since the days of Keynes, economists have taken unemployment as the most important sign of possible under-utilization of labor. However, unemployment is by definition job search, a kind of economic activity of worker, and as such calls for explanation. Besides, unemployment is only a partial indicator of *under-utilization* of labor in the macroeconomy. The celebrated Okun's law which relates the unemployment rate to the growth rate of real GDP demonstrates the significance of under-utilization of employed labor other than unemployment.[1] In this paper, we consider not only unemployment but also on productivity dispersion in the economy.

To consider Keynes' principle of effective demand, we must obviously depart from the Walrasian general equilibrium. The most successful example of "non-Walraian economics" which analyzes labor market in depth is equilibrium search theory surveyed by its pioneers Rogerson et al. (2005), Diamond (2011), Mortensen (2011), and Pissarides (2011). The standard general equilibrium abstracts itself altogether from the search and matching costs which are always present in the actual markets. By explicitly exploring search frictions, search theory has succeeded in shedding much light on the workings of labor market; see also Tobin (1972) for macroeconomics of labor market. While acknowledging the achievement of equilibrium search theory, we find several fundamental problems with the standard theory. In particular, the theory fails to provide a useful framework for explaining *cyclical* changes in effective utilization of labor in the macroeconomy.

Section 2 points out limitations of standard search theory. After brief explanation of the concept of equilibrium based on statistical physics in Sects. 3, 4 presents a

---

[1] Okun (1963) found that a decline of the unemployment rate by one percent raises the growth rate of real GDP by three percent. The Okun coefficient three is much larger than the elasticity of output with respect to labor which is supposed to be equal to the labor share, and roughly one third. This finding demonstrates that there always exists significant under-employment of labor other than unemployment in the macroeconomy: see Okun (1973).

model of stochastic macro-equilibrium. The model explains how the distribution of productivity is determined together with unemployment. Section 5 then explains that the stochastic macro-equilibrium provides a micro-foundation for Keynes' principle of effective demand. It also presents a suggestive evidence supporting the model. The final section offers brief concluding remarks.

2 Limitations of search theory

The search theory starts with the presence of various frictions and accompanying matching costs in market transactions. Once we recognize these problems, we are led to heterogeneity of economic agents and multiple outcomes in equilibrium. In the simplest retail market, for example, with search cost, it would be possible to obtain high and low (more generally multiple) prices for the same good or service in equilibrium. This break with the law of one price is certainly a big step toward reality. Frictions and matching costs are particularly significant in labor market. And the analysis of labor market has direct implications for macroeconomics. In what follows, we discuss labor search theory.

In search equilibrium, potentially similar workers and firms experience different economic outcomes. For example, some workers are employed while others are unemployed. In this way, search theory well recognizes, even emphasizes heterogeneity of workers and firms. Despite this recognition, when it comes to model behavior of economic agent such as worker and firm, it, in effect, presumes the representative agent in the sense that *stochastic economic environment is common to all the agents*; Workers and firms differ only in terms of the realizations of stochastic variable of interest whose probability distribution is common. Specifically, it is routinely assumed that the job arrival rate, the job separation rate, and probability distributions of wages and productivity are common to all the workers and firms.

However, the job separation includes layoffs as well as voluntary quits. It makes no sense that all the firms and workers face the same job separation rate, particularly the probability of layoffs. White collar and blue collar workers face different risks of layoff. The probability of layoffs depends crucially on the state of demand for the firm's product, and as a result, among other things, on industry, region, and ultimately the firm's performance in the product market. The probability is far from being common to all the firms and workers, but is firm and worker-specific.

Similarly, it is difficult to imagine that workers of different educational attainments face the same probability distribution of wage offers; it is plainly unrealistic to assume that a youngster working at gas station faces the same probability of getting a well paid job offer from bank as a graduate of business school. And yet, in standard search models, the assumption that the wage distribution $F(w)$ is common to all the workers is routinely made, and the common $F(w)$ is put into the Bellman equations describing behaviors of firms and workers. This assumption is simply untenable.

Besides, although wages are one of the most important elements in any job offer, workers care not only wages but other factors such as job quality, tenure, and location. Preferences for these other factors which define a job offer certainly differ widely across workers, and are constantly changing over time. Rogerson et al. (2005; p. 962)

say that "although we refer to w as the wage, more generally it could capture some measure of the desirability of the job, depending on benefits, location, prestige, etc." However, this is an illegitimate proposition. All the complexities they refer to simply strengthens the case that we cannot assume that the wage distribution $F(w)$ is common to all the workers. In fact, wage may be even a lexicographically inferior variable to some workers. For example, pregnant female worker might prefer job closer to her home at the expense of lower wage. In summary, workers and firms all act in their own different universes (see Caballero 2010, p. 91).

The second problem of standard search model pertains to the behavior of firm. Though blurred by the standard Poisson modeling, it is routinely assumed that the product market is perfectly competitive in the sense that individual demand curve facing the firm is flat. For example, the flow of revenue generated by employed worker, p is constant and a firm's steady-state profit given the wage offer w is simply $(p-w)l$ where l is the number of workers of this firm. l depends on the firm's wage setting. This is the standard assumption in the literature where p is often called "productivity". Given p, the determination of wages plays a crucial role. It is essentially a model of labor shortage in the sense that the firm's output and profits are determined by labor supply at the level of wages the firm offers to workers. It is curious that the standard search theory makes so much effort to consider the determination of wages within firm taking into account strategic behavior of rival firms while at the same time it leaves the price unexplained under the naïve assumption of perfect competition.

The point is *not* that we must *explicitly* introduce all the complexities characterizing labor and product markets into analytical model. It would simply make model intractable. Rather, we must fully recognize that it is absolutely impossible to trace the microeconomic behaviors, namely decision makings of workers and firms in detail. When the problem is *meso* (say, an oligopolistic market dominated by a few firms, or sales/purchases of old cars in a small town), strategic/optimizing behaviors of economic agents must be explicitly considered. However, the macroeconomy is fundamentally different. In the labor market, microeconomic shocks are indeed *unspecifiable*. Thus, for the purpose of the analysis of the macroeconomy, sophisticated optimization exercises for micro agents based on common probability distribution facing them do not make sense (Aoki and Yoshikawa 2007).

This is actually partly recognized by search theorists themselves. The recognition has led them to introduce the "matching function" into the analysis. The matching function relates the rate of meetings of job seekers and firms to the numbers of the unemployed and job vacancies (Pissarides 2011, pp. 1093–1094). Pissarides recognizes such "real-world features" as differences across workers and jobs; the "universe" differs across workers and firms. Then, at the same time, he recognizes that we need a macro black box. The matching function is certainly a black box not explicitly derived from micro optimization exercises, and is, in fact, not a function of any economic variable which directly affects the decisions of individual workers and firms. Good in spirit, but the matching function is still only a half way in our view.

The matching function is based on a kind of common sense in that the number of job matching would increase when there are a greater number of both job seekers and vacancies. However, it still abstracts itself from an important aspect of reality. As Okun (1973) emphasizes, the problem of unemployment cannot be reduced only to numbers.

The evidence presented above confirms that a high-pressure economy generates not only more jobs than does a slack economy, but also a different pattern of employment. It suggests that, in a weak labor market, a poor job is often the best job available, superior at least to the alternative of no job. A high-pressure economy provides people with a chance to climb ladders to better jobs.

The industry shifts are only one dimension of ladder climbing. Increased upward movements within firms and within industries, and greater geographical flows from lower-income to higher-income regions, are also likely to be significant. (Okun 1973; pp. 234–235)

Dynamics of unemployment cannot be separated from qualities of jobs, or more specifically distribution of productivity on which we focus in the present paper.

To explicitly consider these problems, we face greater complexity and, therefore, need a "greater macro black box" than the standard matching function. Our analysis, in fact, demonstrates that the "matching function" is *not* a structurally given function or technology, but depends crucially on the level of aggregate demand.

3 Stochastic macro-equilibrium: the basic idea

Our vision of the macroeconomy is basically the same as standard search theory. Workers are always interested in better job opportunities, and occasionally change their jobs. While workers search for suitable jobs, firms also search for suitable workers. Firm's job offer is, of course, conditional on its economic performance. The present analysis focuses on the firm's labor productivity. The firm's labor productivity increases thanks to capital accumulation and technical progress or innovations. However, those job sites with high productivity remain only *potential* unless firms face high enough demand for their products; firms may not post job vacancy signs or even discharge the existing workers when demand is low.

We assume that firms are all monopolistically competitive in the sense that they face downward sloping individual demand curves, and the levels of production are determined by demand rather than increasing marginal costs (Solow 1986). Formally, a most elegant general equilibrium model of monopolistic competition is given by Negishi (1960–61). Sweezy (1939) and Negishi (1979) persuasively argue that when the firm is monopolistically competitive, the individual demand curve is actually not only downward-sloping, but must be *kinked at* the current level of output and price. The response of the firm's sales to a change in price is asymmetric because of the asymmetric reactions of rival firms on one hand, and the asymmetric reactions of customers on the other. Drèze (1979) also shows that for a risk-averse firm, uncertainty about the price elasticity of demand has an effect equivalent to that of kinked demand curve with the kink located at the current price and quantity. Therefore, in the economy with uncertainty and frictions as emphasized by search theory, it is reasonable to expect kinked individual demand curve facing monopolistically competitive firm. The corresponding marginal revenue then becomes discontinuous.

Under this assumption, Negishi (1979) shows that a shift in demand is completely absorbed in a change in output leaving the price unchanged because inequality con-

ditions arising from discontinuous marginal revenues remain undisturbed.[2] As Tobin (1993) says, the model opens the door for "quantities determine quantities." In this framework, we can concentrate on quantities (output and labor employment) without explicitly considering price and wages.

Negishi (1979; Chapter 6) shows the existence of general equilibrium of monopolistically competitive firms facing the kinked individual demand curve. Though the equilibrium exists, it is indeterminate because the initial levels of the firms' output and price (the kink point of the individual demand curve) must be exogenously given. To give such initial condition amounts to determining the level of individual demand curve facing each firm. The present analysis provides a rule for allocating the aggregate demand D to micro demand d_k facing each monopolistically competitive firm based on the principle of statistical physics.

The motivation for the method of statistical physics is as follows. Though we assume that firms with higher productivity make more attractive job offers to workers, we do not know how attractive they are to which workers. Whenever possible, workers move to firms with higher productivity, but we never know particular reasons for such moves. For workers to move to firms with higher productivity, it is necessary that those firms must decide to fill the vacant job sites, and post enough number of vacancy signs and/or make enough hiring efforts. They post such vacancy signs and make hiring efforts only when they face an increase of demand for their products, and decide to raise the level of production. It also goes without saying that high productivity firms keep their existing workers only when they face high enough demand.

The question we ask is what the distribution of employed workers is across firms whose productivities differ. As we argued in the pervious section, because microeconomic shocks to both workers and firms are so complex and unspecifiable, optimization exercises based on representative agent assumptions do not help us much. In particular, we never know how the aggregate demand is distributed across monopolistically competitive firms. Besides, among other things, the job arrival rate, the job separation rate, and the probability distribution of wages (or more generally measure of the desirability of the job) differ across workers and firms. This recognition is precisely the starting point of the fundamental method of statistical physics. Foley (1994), in his seminal application of this approach to general equilibrium theory, called the idea "statistical equilibrium theory of markets". Following the lead of Foley, Yoshikawa (2003) applied the concept to macroeconomics. At first, one might think that allowing too large a dispersion of individual characteristics leaves so many degrees of freedom that almost anything can happen. However, it turns out that the methods of statistical physics provide us not only with qualitative results but also with quantitative predictions.

In the present model, the fundamental constraint on the economy as a whole is aggregate demand D. Accordingly, to each firm facing the downward-sloping kinked individual demand curve, the level of demand for its product is the fundamental con-

---

[2] We must note that this result does not ensue only from the assumption that the firm's individual demand curve is downward-sloping. If there is no kink in the perceived demand curve, an increase in sales at the unchanged price may be followed by an increase in price leaving output and labor employment little changed. With the kinked individual demand curve, the marginal revenue becomes discontinuous. Under this assumption, a shift in demand is all absorbed in changes in output and labor employment leaving the price unchanged.

straint. The problem is how the aggregate demand D is allocated to these monopolistically competitive firms, namely the determination of d_k.

$$D = \sum_{k=1}^{K} d_k \qquad (1)$$

Our model provides a solution. The basic idea behind the analysis can be explained with the help of the simplest case. For the moment, we focus on productivity dispersion. We will introduce unemployment into the model in the next section.

Suppose that n_k workers belong to firms whose productivity is c_k ($c_k < c_{k'}$ where $k < k'$). There are K levels of productivity in the economy ($k = 1, 2, \ldots, K$). The total number of workers N is given.

$$\sum_{k=1}^{K} n_k = N \qquad (2)$$

A vector $n = (n_1, n_2, \ldots, n_K)$ shows a particular allocation of workers across firms with different productivities. The combinatorial number W_n of obtaining this allocation, n, is equal to that of throwing N balls to K different boxes. Because the number of all the possible ways to allocate N different balls to K different boxes is K^N, the probability that a particular allocation $n = (n_1, n_2, \ldots, n_K)$ is obtained is

$$P_n = \frac{W_n}{K^N} = \frac{1}{K^N} \frac{N!}{\prod_{k=1}^{K} n_k!}. \qquad (3)$$

It is the fundamental postulate of statistical physics that the state or the allocation $n = (n_1, n_2, \ldots, n_K)$ which maximizes the probability P_n or (3) *under macro-constraints* is to be realized.[3] The idea is similar to maximum likelihood in statistics/econometrics. Maximizing P_n is equivalent to maximizing $\ln P_n$. Applying the Stirling formula for large number, we find that the maximization of $\ln P_n$ is equivalent to that of S.

$$S = -\sum_{k=1}^{K} p_k \ln p_k \quad (p_k = \frac{n_k}{N}) \qquad (4)$$

S is the Shannon entropy, and captures the combinatorial aspect of the problem. Though the combinatorial consideration summarized in the entropy plays a decisive role for the final outcome, that is not the whole story, of course. The qualification "under macro-constraints" is crucial.

The first macro-constraint concerns the labor endowment, (2). The second macro-constraint concerns the effective demand. Because firm's production $c_k n_k$ is constrained by demand d_k, we obtain

---

[3] To be precise, it is to be realized in the sense of expected value. In physics, variance is normally so small relative to expected value that we practically always observe the expected value.

$$D = \sum_{k=1}^{K} d_k = \sum_{k=1}^{K} c_k n_k \tag{5}$$

Here, aggregate demand D is assumed to be given. In our analysis, we explicitly analyze the allocation of labor $(n_1, n_2, \ldots, n_K)$. The allocation of labor basically corresponds to the allocation of the aggregate demand to monopolistically competitive firms.

To maximize entropy S *under two macro-constraints* (2) and (5), set up the following Lagrangean form L:

$$L = -\sum_{k=1}^{K} \left(\frac{n_k}{N}\right) \ln \left(\frac{n_k}{N}\right) + \alpha \left[N - \sum_{k=1}^{K} n_k\right] + \beta \left[D - \sum_{k=1}^{K} c_k n_k\right] \tag{6}$$

with two Lagrangean multipliers, α and β. Maximization of this Lagrangean form with respect to n_k leads us to the first-order conditions:

$$\ln \left(\frac{n_k}{N}\right) = -1 - \alpha N - \beta N c_k \quad (k = 1, 2, \ldots, K). \tag{7}$$

Because n_k/N sums up to one, we obtain

$$\frac{n_k}{N} = \frac{e^{-\beta N c_k}}{\sum_{k=1}^{K} e^{-\beta N c_k}} \tag{8}$$

Thus, the number of workers working at firms with productivity c_k is exponentially distributed. It is known as the Boltzmann distribution in physics.

Here arises a crucial difference between economics and physics. In physics, c_k corresponds to the level of energy. Whenever possible, particles tend to move toward the *lowest* energy level. To the contrary, in economics, workers always strive for better jobs offered by firms with higher productivity c_k. As a result of optimization under unobservable respective constraints, workers move to better jobs. In fact, if allowed, all the workers would move up to the job sites with the highest productivity, c_K. This situation corresponds to the textbook Pareto optimal Walrasian equilibrium with no frictions and uncertainty. However, this state is actually impossible unless the level of aggregate demand D is so high as equal to the maximum level $D_{\max} = c_K N$. When D is lower than $D_{\max}$, the story is quite different. Some workers — a majority of workers, in fact, must work at job sites with productivity lower than c_K.

How are workers distributed over job sites with different productivity? Obviously, it depends on the level of aggregate demand. When D reaches its lowest level, $D_{\min}$, workers are distributed evenly across all the sectors with different levels of productivity, $c_1, c_2, \ldots, c_K$. Here, $D_{\min}$ is defined as $D_{\min} = N(c_1 + c_2 + \cdots + c_k)/K$. It is easy to see that the lower the level of D is, the greater the combinatorial number of distribution $(n_1, n_2, \ldots, n_K)$ which satisfies aggregate demand constraint (5) becomes.

As explained above, the combinatorial number W_n of a particular allocation $n = (n_1, n_2, \ldots, n_K)$ is basically equivalent to the Shannon entropy, S defined by (4). S

Fig. 1 Entropy S and aggregate demand D. *Note* β is a Lagrangean multiplier in Eq. (6) in the text
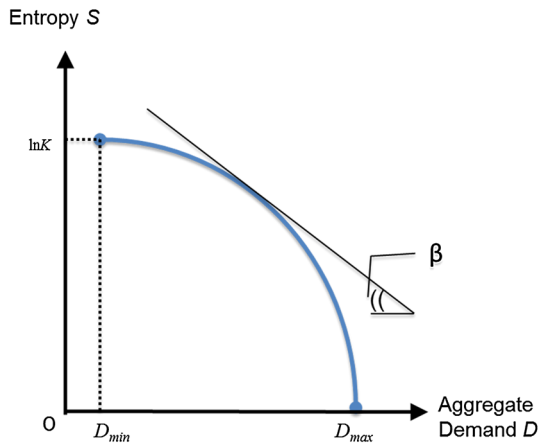
increases when D decreases. For example, in the extreme case where D is equal to the maximum level $D_{\max}$, all the workers work at job sites with the highest productivity. In this case, the entropy S becomes zero, its lowest level because $n_K/N = 1$ and $n_k/N = 0$ $(k \neq K)$. In the other extreme where aggregate demand is equal to the minimum level $D_{\min}$, we have $n_K = N/K$, and the entropy S defined by (4) becomes $\ln K$, its maximum level. The relation between the entropy S and the level of aggregate demand D, therefore, looks like the one shown in Fig. 1.

At this stage, we can recall that the Lagrangean multiplier β in (6) for aggregate demand constraint is equal to

$$\beta = \frac{\partial L}{\partial D} = \frac{\partial S}{\partial D}. \tag{9}$$

β is the slope of the tangent of the curve as shown in Fig. 1, and, therefore, is negative.[4] With negative β, the exponential distribution (8) is upward-sloping. However, unless the aggregate demand is equal to (or greater than) the maximum level, $D_{\max}$, workers' efforts to reach job sites with the highest productivity c_K must be frustrated because firms with the highest productivity do not employ a large number of workers and are less aggressive in recruitment, and accordingly, it becomes harder for workers to find such jobs. As a consequence, workers are distributed over all the job-sites with different levels of productivity.

The maximization of entropy under the aggregate demand constraint (6), in fact, balances two forces. On one hand, whenever possible, workers move to better jobs identified with job sites with higher productivity. It is the outcome of successful job

---

[4] In physics, β is normally positive. This difference arises because workers strive for job sites with higher productivity, not the other way round (Iyetomi 2012). In physics, β is equal to the inverse of temperature, or more precisely, temperature is defined as the inverse of $\partial S/\partial D$ when S is the entropy and D energy. Thus, negative β means the negative temperature. It may sound odd, but the notion of negative temperature is perfectly legitimate in such systems as the one in the present analysis; see Appendix E of Kittel and Kroemer (1980).

matching resulting from the worker's search and the firm's recruitment. When the level of aggregate demand is high, this force dominates. However, when D is lower than $D_{\max}$, there are in general a number of different allocations $(n_1, n_2, \ldots, n_K)$ which are consistent with D.

As we explained in the previous section, micro shocks facing both workers and firms are truly unspecifiable. We simply do not know which firms with what productivity face how much demand constraint and need to employ how many workers with what qualifications. We do not know which workers are seeking what kind of jobs with how much productivity, either. Here comes the maximization of entropy. It gives us the distribution $(n_1, n_2, \ldots, n_K)$ which corresponds to the maximum combinatorial number consistent with given D.

The entropy maximization under aggregate demand constraint plays, therefore, the role similar to the matching function in the standard search theory. Note that unlike the standard matching function which focuses only on the number of jobs, the matching of job quality characterized by productivity plays a central role in the present analysis. The matching of high productivity jobs is ultimately conditioned by the level of aggregate demand. That is, uncertainty and frictions emphasized by the standard search theory are *not* exogenously given, but depend crucially on aggregate demand. In a booming gold-rush town, one does not waste a minute to find a good job! The opposite holds in a depressed city.

It is essential to understand that the present approach does *not* regard economic agents' behaviors as random. Certainly, firms and workers maximize their profits and utilities. The present analysis, in fact, presumes that workers always strive for better jobs characterized by higher productivity. Randomness underneath the entropy maximization comes from the fact that both the objective functions of and constraints facing a large number of economic agents are constantly subject to *unspecifiable* micro shocks. We must recall that the number of households is of order 10^7, and the number of firms, 10^6. Therefore, there is nothing for outside observers, namely economists analyzing the macroeconomy but to regard a particular allocation *under macro-constraints* as equi-probable. Then, it is most likely that the allocation of the aggregate demand and workers which maximizes the probability P_n or (3) *under macro-constraints* is realized.[5]

---

[5] This method has been time and again successful in natural sciences when we analyze object comprising many micro elements. Economists might be still skeptical of the validity of the method in economics saying that inorganic atoms and molecules comprising gas are essentially different from optimizing economic agents. Every student of economics knows that behavior of dynamically optimizing economic agent such as the Ramsey consumer is described by the Euler equation for a problem of calculus of variation. On the surface, such a sophisticated economic behavior must look remote from "mechanical" movements of an inorganic particle which only satisfy the law of motion. However, every student of physics knows that the Newtonian law of motion is actually nothing but the Euler equation for a certain variational problem; particles minimize the energy or the Hamiltonian!. It is called the principle of least action: see Chapter 19 of Feynman (1964)'s *Lectures on Physics*, Vol. II. Therefore, behavior of dynamically optimizing economic agent and motions of inorganic particle are on a par to the extent that they both satisfy the Euler equation for respective variational problem. The method of statistical physics can be usefully applied not because motions of micro units are "mechanical," but because object or system under investigation comprises many micro units individual movements of which we are unable to know.

Fig. 2 Distribution of labor productivity in Japan (2007)

Mean= 1.29×10^4, Standard Deviation= 2.37×10^4
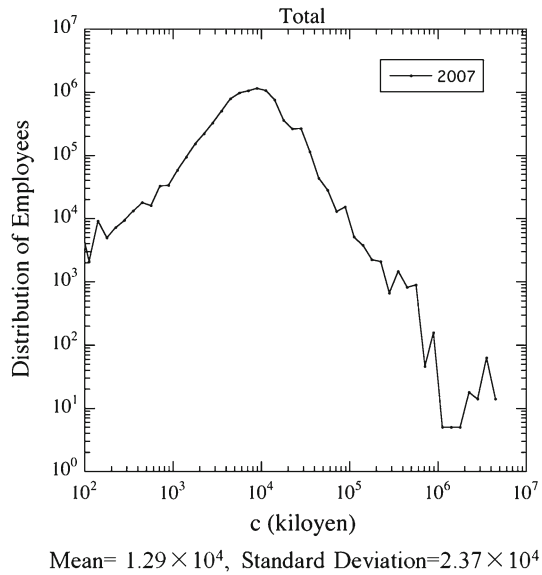
4 The model

The above analysis shows that the distribution of workers at firms with different productivities depends crucially on the level of aggregate demand. Though the simple model is useful to explain the basic idea, it is too simple to apply to the empirically observed distribution of labor productivity. Besides, it abstracts itself from unemployment. We introduce unemployment into the model.

4.1 Empirical distribution of productivity

Figure 2 shows the distribution of workers at different productivity levels for the Japanese economy. The data used are the Nikkei Economic Electric Database (NEEDS, http://www.crd-office.net/CRD/english/index.html) and the Credit Risk Database (CRD, http://www.crd-office.net/CRD/english/index.html) which cover more than a million large and medium/small firms for 2007.

The "productivity" here is simply value added of the firm divided by the number of employed workers, that is, the average labor productivity. Theoretically, we should be interested in unobserved *marginal* productivity, not the average productivity. Besides, proper "labor input" must be in terms of work hour, or for that matter even in terms of work efficiency units rather than the number of workers. For these reasons, the average labor productivity shown in the figure is a crude measure of theoretically meaningful unobserved marginal productivity. However, Aoyama et al. (2010; pp. 38–41) demonstrates that when the average productivity and measurement errors are independent, the distribution of true marginal productivity obeys the power law with the same exponent as that for the measured average productivity. In other words, distribution is robust with respect to measurement errors in the present case.

Figure 2 drawn on the double logarithm plane broadly shows that (1) the distribution of labor productivity is single-peaked, (2) in the low productivity (left) region, it is upward-sloping exponential whereas (3) in the high productivity (right) region, it obeys downward-sloping power-law (Aoyama et al. 2010). Ikeda and Souma (2009) find a similar distribution of productivity for the US while Delli Gatti et al. (2008) find power-law tails of productivity distribution for France and Italy. In what follows, we present an extended model for explaining the broad shape of this empirically observed distribution. The extended model also explains unemployment.

We explained in the previous section that the entropy maximization under macro constraints leads us to exponential distribution. This distribution with negative β can explain the broad pattern of the left-hand side of the distribution shown in Fig. 2, namely an upward-sloping exponential distribution (Iyetomi 2012). However, we cannot explain the downward-sloping power distribution for high productivity firms. To explain it, we need to make an additional assumption that the number of potentially available high-productivity jobs is limited, and that it decreases as the level of productivity rises.

Potential jobs f_j are created by firms by accumulating capital and/or introducing new technologies, particularly new products. On the other hand, they are destroyed by firms' losing demand for their products permanently. Schumpeterian innovations by way of creative destruction raise the levels of some potential jobs, but at the same time lower the levels of others. In this way, the number of potential jobs with a particular level of productivity keeps changing. Note, however, that they remain only *potential* because firms do not necessarily attempt to fill all the job sites with workers. To fill them, firms either keep the existing workers on the job or post job vacancy signs and make enough hiring efforts, but they are economic decisions, and depend crucially on the economic conditions facing firms. The number of potential job sites, therefore, is not exactly equal to, but rather imposes a ceiling on the sum of the number of filled job sites, or employment and the unfilled jobs.

4.2 Distribution of productivity

Under reasonable assumptions, distribution of potential job sites with high productivity becomes downward-sloping power law. Adapting the model of Marsil and Zhang (1998), we can derive a power-law distribution such as the one for the tail of the empirically observed distribution of labor productivity; see Yoshikawa (2013) for details. However, the determination of employment by firms with various levels of productivity is another matter. To fill potential job sites with workers is the firm's economic decision. The most important constraining factor is the level of demand facing the firm in the product market. To fill potential job sites, the firm must either keep the existing workers on the job, or make enough hiring efforts including posting vacancy signs toward successful job matching. Such actions of the firms and job search of workers are purposeful. However, micro shocks affecting firms and workers are just unspecifiable. Then, how are workers actually employed at firms with various levels of productivity? This is the problem we considered in the previous section. In what follows, we will consider it in a more general framework.

The number of workers working at the firms with productivity, c_k, namely n_k is

$$n_k \in \{0, 1, \ldots, f_k\} \quad (k = 1, 2, \ldots, K). \tag{10}$$

Here, f_k is the number of potential jobs with productivity c_k, and puts a ceiling on n_k.[6] We assume that in the low productivity region, f_k is large enough meaning that n_k is virtually unconstrained by f_k. In contrast, in the high productivity region, f_k constrains, n_k and its distribution is power distribution as we have analyzed above.[7]

The total number of employed workers is simply the sum of n_j:

$$N = \sum_{k=1}^{K} n_k. \tag{11}$$

In the basic model in Sect. 3, the total number of employed workers, N is exogenously given (Eq. 2). In the extended model, N is assumed to be variable. N is smaller than the exogenously given total number of workers or labor force, $L(N < L)$. The difference between L and N is the number of the unemployed, U:

$$U = L - N. \tag{12}$$

As in the basic model, firms are monopolistically competitive facing the downward-sloping kinked individual demand curve. The firm's output is constrained by demand, which is conditioned by the level of aggregate demand, D (Eq. 5). In the basic model, D is literally given. Accordingly, total output is also constant. In the present model, we more realistically assume that in accordance with fluctuations of aggregate demand, total output Y also fluctuates. Specifically, Y defined by

$$Y = \sum_{k=1}^{K} y_k = \sum_{k=1}^{K} c_k n_k \tag{13}$$

---

[6] When the number of potential jobs with high productivity is limited, behavior of economic agents necessarily becomes correlated; If good jobs are taken by some workers, it becomes more difficult for others to find such jobs. Galibaldi and Scalas (2010) suggest that we study the problem by Markov model with such constraints. The present analysis precisely does it by introducing ceilings on n_j.

[7] Like many others, Postel-Vinay and Robin (2002) assume that there is no ceiling for job-sites with high productivity. On this assumption as well as the assumption that the product market is perfectly competitive, they regard a decreasing number of workers with high productivity jobs as a consequence of less recruitment efforts made by high productivity firms than by low productivity firms. This is a strange interpretation. There is no reasonable reason why high productivity firms make less recruitment efforts than low productivity firms. A more plausible assumption is that firms are monopolistically competitive facing the downward-sloping individual demand curve rather than the price takers in the product market, and that jobs with high productivity are limited in number by such demand constraints. Suppose, for example, that the automobile industry is a high productivity industry. It would be unreasonable to argue that the level of employment in the industry is only the outcome of job matching, and that a limited size of employment is due to a lack of the firms' recruitment efforts. Plainly, the size of car producers is given by their capacity which is, in turn, determined by the level of demand for cars.

is now stochastic, and its expected value $\langle Y \rangle$ is equal to constant D. That is, we have

$$\langle Y \rangle = D. \tag{14}$$

Aggregate demand constrains total output in the sense of its expected value.

Under this assumption, the probability of total output Y turns out to be exponential;[8] The density function $g(Y)$ is

$$g(Y) = \frac{e^{-\beta Y_i}}{\sum_i e^{-\beta Y_i}} \tag{15}$$

Obviously, Y constrained by aggregate demand D affects the distribution of workers, n_k (Eq. 13). In the present model, the number of employed workers N is not constant, but changes causing changes in unemployment. Besides, the number of potential job sites with high productivity, f_j constrains n_j. Under these assumptions, we seek the state which maximizes the probability P_n or Eq. (3).

Before we proceed, it is necessary to explain *partition function* because it is rarely used in economics, but we will use it intensively in the subsequent analysis. When a stochastic variable Y is exponentially distributed, that is, its density function $g(Y)$ is given by Eq. (15), partition function Z is defined as

$$Z = \sum_i e^{-\beta Y_i}. \tag{16}$$

This function is extremely useful as moment generating function. For example, the first moment or the average of Y can be simply found by differentiating $\log Z$ with respect to β.

$$-\frac{d \log Z}{d\beta} = -\frac{d}{d\beta} \log \left(\sum_i e^{-\beta Y_i} \right) = -\frac{\sum_i (-Y_i) e^{-\beta Y_i}}{\sum_i e^{-\beta Y_i}}$$

$$= \sum_i Y_i \left(\frac{e^{-\beta Y_i}}{\sum_i e^{-\beta Y_i}} \right) = \sum_i Y_i g(Y_i) = E(Y_i) \tag{17}$$

As in the basic model, we want to find the state which maximizes the probability, P_n or Eq. (3). We have two macro-constraints, Eqs. (11) and (13). The total number of workers employed N is, however, not constant but variable. The aggregate output Y is also not constant but obeys the exponential distribution, namely Eq. (15).

---

[8] This result is obtained by the method of Gibbs' canonical ensemble. Gibbs established the statistical mechanics by introducing the concept of "canonical ensemble" which is a collection of macro states, Y in our present case. Suppose that there are K possible levels of Y denoted by $Y_1, \ldots, Y_K$. For the moment, we reinterpret n_k as the number of cases where Y takes the value Y_k ($k = 1, \ldots, K$). The sum of n_k, N is given. Then, n_k satisfies Eq. (2). We assume that the average of Y is equal to constant D: $\sum_{k=1}^{K} Y_k (n_k/N) = D$. Replacing c_k by Y_k/N, we observe that this equation is equivalent to Eq. (5). Thus, we can apply the exactly same entropy maximization as we did in the basic model in Sect. 3. It leads us to Eq. (15).

Because the level of total output depends on the total number of employed workers N, we denote Y_i as $Y_i(N)$. Then, the canonical partition function Z_N can be written as

$$Z_N = \sum_i e^{-\beta Y_i(N)}. \tag{18}$$

Using Eq. (13), we can rewrite this partition function as follows:

$$Z_N = \sum_{\{n_k\}} \exp\left(-\beta \sum_{k=1}^{K} c_k n_k\right). \tag{19}$$

It is generally difficult to carry out the summation with respect to $\{n_k\}$ under constraint (11), namely $N = \sum n_k$. Rather than taking N as given, we better allow N to be variable as we do here, and consider the following extended partition function Φ.[9] With the Lagrangean multiplier for N, μ, this partition function is defined as

$$\Phi = \sum_{N=0}^{\infty} z^N Z_N \tag{20}$$

where

$$z = e^{\beta\mu}. \tag{21}$$

As the Lagrangean multiplier for N, the parameter μ is the marginal contribution of an addition employment to the entropy of the macroeconomy.[10] Because the entropy corresponds one to one to the level of aggregate demand as shown in Fig. 1, it measures the marginal product of a worker who newly acquired job out of the pool of unemployment. Thus, μ plays a role similar to the reservation wage in standard models. When μ is high, the unemployed worker is "choosy", and *vice versa*. In this sense, μ is a barrier between unemployment and employment. However, there is a fundamental difference between the reservation wage and μ. In standard models, the reservation wage R is absolute in the sense that no worker is willing to work with the wages below R. This concept is useful only if R can be well defined and known. As we pointed out in Sect. 2, the equilibrium search theory effectively presumes the representative worker so that the reservation wage R can be uniquely defined.

In contrast, μ in the present analysis plays only a relative role as a barrier between unemployment and employment. Each worker has his/her own "reservation job attractiveness" which we never observe. μ represents the average reservation job attractiveness of a heterogeneous group of unemployed workers. We will explain the meaning of

---

[9] It is called the grand canonical partition function in physics.

[10] μ is called the chemical potential in physics, and measures the marginal contribution in terms of energy of an additional particle to the system under investigation.

μ later. For the moment, it is important to note that depending on the level of aggregate demand, some workers accept jobs with productivity below μ.

Substituting Eq. (19) into Eq. (20), the grand canonical partition function Φ becomes as follows:

$$\Phi = \sum_{N=0}^{\infty} z^N \sum_{nj} \exp\left\{-\beta \sum_j n_j c_j\right\} \quad \text{where } z = e^{\beta\mu} \tag{22}$$

Using the definitions of z, (21), and also N, (11), we have

$$\Phi = \sum_{N=0}^{\infty} e^{\beta\mu(n_1+\cdots+n_K)} \sum_{n_j} \exp\left\{-\beta \sum_j n_j c_j\right\} = \prod_{j=1}^{K} \sum_{nj} \exp[\beta(\mu - c_j)n_j]. \tag{23}$$

Because there is ceiling f_j for n_j [constraint (10)], (23) can be rewritten as follows:

$$\Phi = \prod_{j=1}^{K} \left[1 + e^{\beta(\mu-c_j)} + \cdots e^{f_j\beta(\mu-c_j)}\right] = \prod_{j=1}^{K} \left[\frac{1 - e^{(f_j+1)\beta(\mu-c_j)}}{1 - e^{\beta(\mu-c_j)}}\right] \tag{24}$$

With this grand canonical partition function Φ, we can easily obtain the expected value of the total number of employed workers N, $\langle N \rangle$ by differentiating $\log \Phi$ with respect to μ which corresponds to the reservation wage of the unemployed worker. This can be seen by differentiating (20) and noting the definition of z (21).

$$\frac{1}{\beta}\left[\frac{\partial}{\partial\mu} \log \Phi\right] = \frac{1}{\beta}\left[\frac{\partial}{\partial\mu} \log\left(\sum_{N=0}^{\infty} e^{\beta\mu N} Z_N\right)\right]$$

$$= \frac{1}{\beta}\left[\frac{\beta \sum_{N=0}^{\infty} N e^{\beta\mu N} Z_N}{\sum_{N=0}^{\infty} e^{\beta\mu N} Z_N}\right] = \langle N \rangle. \tag{25}$$

In the present case, Φ is actually given by equation (24). Therefore, we can find $\langle N \rangle$ as follows.

$$\langle N \rangle = \frac{1}{\beta}\left[\frac{\partial}{\partial\mu} \log \Phi\right]$$

$$= \frac{1}{\beta} \sum_{j=1}^{K} \frac{\partial}{\partial\mu} \{\log(1 - e^{(f_j+1)\beta(\mu-c_j)}) - \log(1 - e^{\beta(\mu-c_j)})\}$$

$$= \sum_{j=1}^{K} \left[\frac{(f_j + 1)e^{(f_j+1)\beta(\mu-c_j)}}{e^{(f_j+1)\beta(\mu-c_j)} - 1} - \frac{e^{\beta(\mu-c_j)}}{e^{\beta(\mu-c_j)} - 1}\right] \tag{26}$$
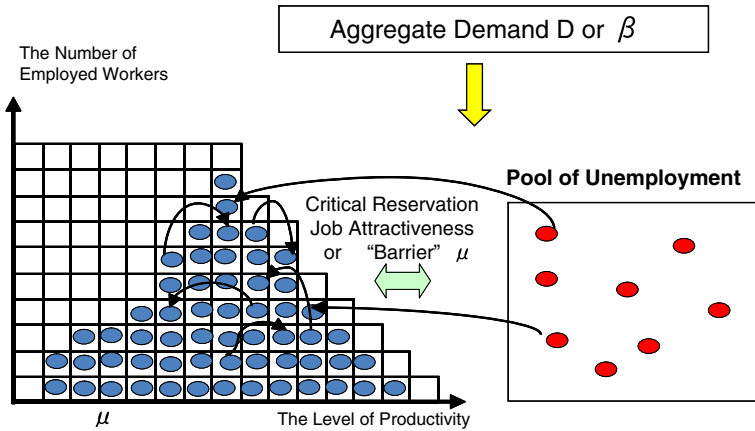
Fig. 3 Model of stochastic macro-equilibrium

The expected value of the number of workers employed on the job sites with productivity c_j, $\langle n_j \rangle$ is simply the corresponding term in the summation of $< N >$ or Eq. (26).

$$\langle n_j \rangle = \left[\frac{(f_j + 1)e^{(f_j+1)\beta(\mu-c_j)}}{e^{(f_j+1)\beta(\mu-c_j)} - 1} \right] - \left[\frac{e^{\beta(\mu-c_j)}}{e^{\beta(\mu-c_j)} - 1} \right] \tag{27}$$

Equation (27) determines the distribution of workers across job-sites with different levels of productivity in our stochastic macro-equilibrium. Figure 3 shows how this model works. Most important, the distribution is fundamentally conditioned by aggregate demand. When the level of aggregate demand is high, it is more likely that high productivity firms employ more workers. They keep the existing workforce, and attract not only the unemployed but also workers currently on the inferior jobs. As Okun (1973) vividly illustrates, "a high pressure economy provides people with a chance to climb ladders to better jobs." And people actually climb ladders in such circumstances.

4.3 Unemployment and the reservation job attractiveness

The distribution of employed workers across job-sites with different productivity simultaneously determines the level of unemployment because the number of unemployed U is simply the difference between exogenously given labor force, L and the number of employed workers, N given by (26). When aggregate demand D rises, more workers work at job sites with higher productivity while at the same time, unemployment decreases.

An increase in μ *ceteris paribus* raises unemployment. This can be most clearly seen in the case of the extremely high aggregate demand. In such a case ($\beta = -\infty$), μ plays a role similar to the reservation wage in the standard theory in the sense that no worker takes a job whose productivity is lower than μ. Thus, it is trivial to see that an increase in μ lowers employment and raises unemployment. More generally, however,

we must recognize that the reservation job attractiveness differs across workers, and that it cannot be reduced to a single number of a particular variable such as wages. μ in this model is a "barrier" between unemployment and employment, but unlike the reservation wage, it is not an absolute barrier. Given μ, when aggregate demand is low, workers actually work at job-sites whose productivity is lower than μ. As we will see shortly, the effect of a change in μ on unemployment depends on the level of aggregate demand.

Unemployment which arises because of frictions and uncertainty in the labor market is by definition job search. Each worker has his/her own reservation job attractiveness. However, frictions and uncertainty depend on the level of aggregate demand. When aggregate demand is high, high productivity firms make many job openings so that workers can find attractive jobs easily. When aggregate demand is low, just the opposite holds true. In this sense, some of the "frictional" unemployment is actually caused by aggregate demand deficiency. Thus, *the definition of frictional unemployment, let alone natural unemployment* (Friedman 1968), *must be necessarily ambiguous unless the level of aggregate demand is extremely high.*

4.4 A numerical example

With the help of a simple numerical example, we can better understand how Eq. (27) looks like, and also how unemployment is determined in the present model. In this example, the level of productivity are assumed to be $c_1 = 1, \ldots, c_{200} = 200$. μ is set equal to 25. The labor force L is assumed to be 630. The number of potential jobs or the ceiling at each productivity level, f_j, is assumed to be constant at 10 for $c_1, \ldots, c_{50}$, while it declines for c_j ($j = 50, \ldots, 200$) as c_j increases. Specifically, for c_j ($j = 50, \ldots, 200$), f_j obeys a power distribution: $f_j \sim 1/c_j^2$. This assumption means that low productivity jobs are potentially abundant whereas high productivity jobs are limited.

Figure 4 shows two cases; Case A corresponds to high aggregate demand whereas Case B to low aggregate demand. Specifically, β is assumed to be (A) -0.05, and (B) -0.02. Case (A) $\beta = -0.05$ corresponds to high demand D whereas Case (B)$\beta = -0.02$ to low demand. In Fig. 4, we observe that n_j increases up to $j = 50$, and then declines from $j = 50$ to 200 in both cases. Whenever possible, workers strive to get better jobs offered by firms with higher productivity. That is why the number of workers n_j increases as the level of productivity rises in the relatively low productivity region. Note that the number of potential jobs or the ceiling is constant in this region, and yet that n_j increases as the level of productivity c_j rises. The number of workers n_j turns to be a decreasing function of productivity c_j in the high productivity region simply because the number of potentially available jobs f_j declines as c_j rises. Note, however, that n_j is not equal to f_j which is shown by a dotted line in the figure; n_j is strictly smaller than f_j. Single-peaked distribution shown in Fig. 4 is broadly consistent with the observed pattern of productivity dispersion (Fig. 2).

When aggregate demand increases, the number of employed workers N which corresponds to the area below the distribution curve, increases. Specifically, N is 618 in case (A) while it is 582 in case (B). It means that given labor force $L = 630$, the
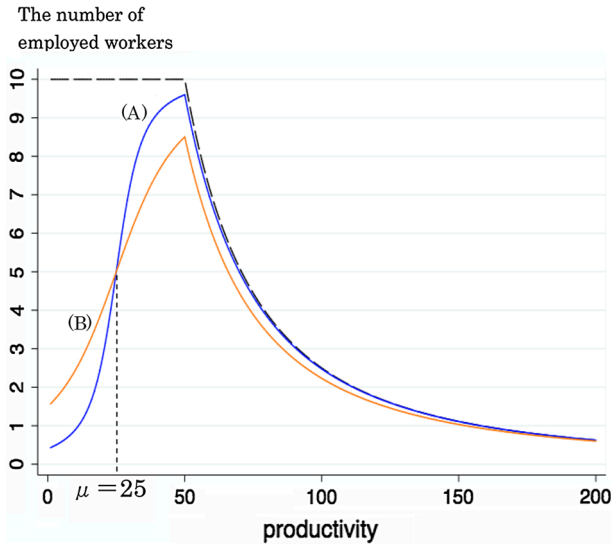
The number of
employed workers

Fig. 4 Distribution of Productivity and Aggregate Demand. *Note* (*A*) High aggregate demand ($\beta = -0.05$), (*B*) low aggregate demand ($\beta = -0.02$). See the main text for details

Table 1 Aggregate demand, job quality, and unemployment

Aggregate demand β	← High			Low →		
	−0.1	−0.05	−0.03	−0.02	−0.01	−0.001
Unemployment rate (U/L) %	0.9	1.8	4.1	7.6	15.9	29.2
Share of employment with productivity $c_j > \mu$ (%)	97	94	91	88	83	74
Share of employment with productivity $c_j > \mu$ (%)	3	6	9	12	17	26

μ = The critical "reservation job attractiveness"
See the main text for details

unemployment rate $U/L = (L - N)/L$ declines when aggregate demand D rises. In this example, the unemployment rate is 1.8 % in case (A) while it is 7.6 % in case (B). Table 1 shows (a) the unemployment rate, (b) the share of workers on job sites with productivity higher than μ, and (c) the corresponding share of workers on job sites with productivity lower than μ for various levels of aggregate demand or β. When aggregate demand rises, the unemployment rate declines while at the same time, the share of high quality job goes up.

Using this numerical example, we can also consider the effects of a change in the reservation job attractiveness μ on employment / unemployment. In the base model, μ is set equal to 25. Table 2 shows how employment and unemployment change when μ increases from 25 to 28. μ corresponds to the reservation wage in standard models so that an increase in μ raises the unemployment rate. However, Table 2 shows that the extent to which the unemployment rate rises depends crucially on the level of aggregate demand. The higher the level of aggregate demand is, the more significant

Table 2 The effects of a change in the reservation job attractiveness μ on unemployment depends on the level of aggregate demand

Aggregate demand β	← Low			High →	
	−0.001	−0.005	−0.01	−0.02	−0.05
(1) Employment when $\mu = 25$	446	487	530	582	618
(2) Employment when $\mu = 28$	444	477	513	557	590
(3) Change in employment when μ increased from 25 to 28 [(2)–(1)]	−2	−10	−17	−25	−28
(4) Change in unemployment rate ((3)/$L = 630$)	0.3	1.6	2.7	4.0	4.5

the effect of an increase μ on the unemployment rate is. Thus, the effects of a change in μ on unemployment gets smaller when the level of aggregate demand is low; in deep recession, the "reservation wage" does not play a major role as a determinant of unemployment.

5 The principle of effective demand

Keynes (1936) argued that the aggregate demand determines the level of output in the economy as a whole. Factor endowment and technology may set a ceiling on aggregate output, but the actual level of output is effectively determined by aggregate demand. Unemployment is obviously a problem of the labor market, but Keynes argued that it is basically caused by demand deficiency in the market for goods and services; Changes in aggregate demand determine cyclical changes in unemployment.

Our economy really experiences occasional aggregate demand shocks (See Iyetomi et al. 2011 for details). The post-Lehman "great recession" provides us with an excellent example of the negative aggregate demand shock. The unemployment rate certainly rose. The relatively low and cyclically insensitive Japan's unemployment rate was 3.6 % as of July 2007, but after the global financial crisis, rose to 5.5 percent by July 2009. During the same period, the unemployment rate rose from 4.6 to 9.8 % in the US.

However, unemployment is not the whole story because the distribution of productivity also changes. Figure 5a–c compares the distributions of productivity *before* and *after* the Lehman crisis, namely 2007 and 2009; (a) total, (b) manufacturing sector, and (c) non-manufacturing sector. As our theory indicates, the distribution as a whole, in fact, tilts toward lower productivity in sever recession. Figure 5 shows that the tilt of the distribution toward low productivity is more conspicuous for the manufacturing industry than for the non-manufacturing industry. It is due to the fact that in Japan, the 2009 recession after the bankruptcy of the Lehman Brothers was basically caused by a fall of exports, and that exports consist mainly of manufactured products such as cars. We can observe, however, that the distribution tilts toward low productivity for the non-manufacturing industry as well, particularly in the high productivity region. This is, of course, due to the fact that a fall of demand in the manufacturing sector spills over to the non-manufacturing sector.

Fig. 5 Distributions of labor productivity in Japan before (2007) and after (2009) the Lehman crisis
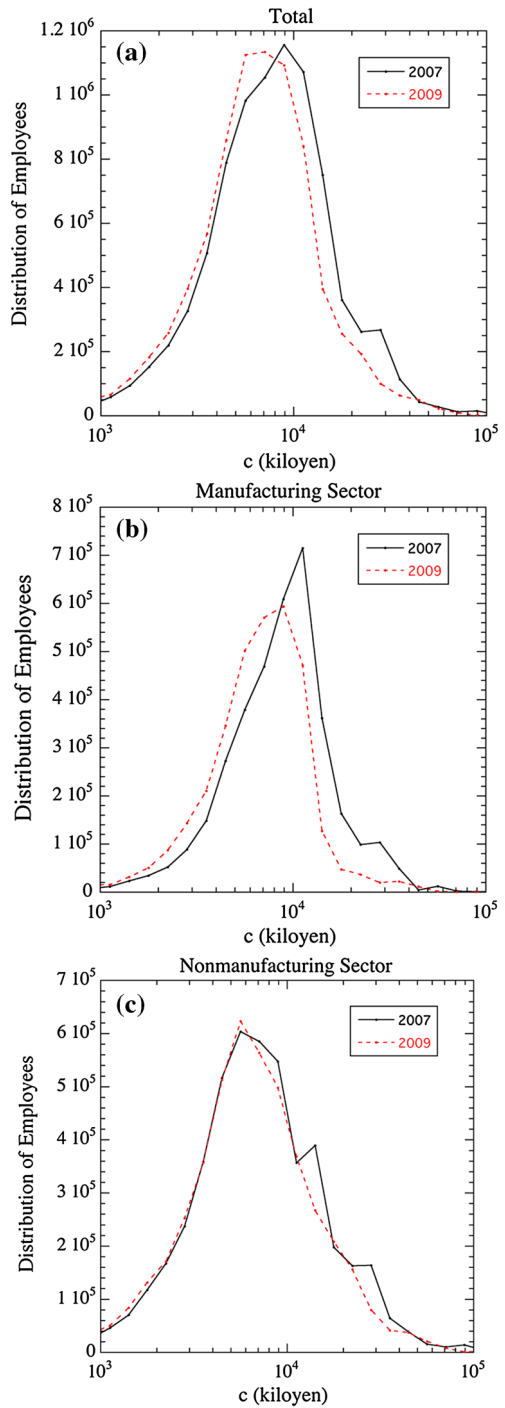
Following the lead of Kydland and Prescott (1982) the standard literature in macro-economics focuses on "productivity shocks," and takes them as exogenous. Our analysis shows, however, that cyclical changes in "productivity" can be nothing but the *result* of aggregate demand shocks. When output changes responding to changes in aggregate demand, the level of utilization of production factors must necessarily change. An example is cyclical changes of capacity utilization of capital. Unemployment of labor is another. However, as a measure of underutilization of labor, unemployment is not sufficient because distribution of productivity of employed workers also significantly changes. Our model shows that aggregate demand determines the degree of uncertainty and frictions with respect to dispersion of job quality and the rate of successful matching in the labor market. As a consequence, both unemployment and distribution of labor productivity across firms and job sites change. This is the market mechanism beneath Keynes' principle of efficient demand.

6 Concluding remarks

It is a cliché that the Keynesian problem of unemployment and under-utilization of production factors arises because prices and wages are inflexible. Tobin (1993) as well as Keynes (1936) himself argued that the principle of effective demand holds true regardless of flexibility of prices and wages. The essential point is that adjustment in terms of quantity is faster, than that of price.

The natural micro picture underneath the Keynesian economics is monopolistic competition of firms facing the downward sloping individual demand curve, certainly not perfect competition in the product market (Solow 1986). Negishi (1979)'s model of general equilibrium of monopolistically competitive firms with the kinked individual demand curve provides a neat microfoundation for what Tobin (1993) called the Old Keynesian view in which "quantities determine quantities." This model, however, abstracts itself from frictions and uncertainty present in the labor market.

The standard equilibrium search theory has filled a gap by explicitly considering frictions and matching costs in the labor market. While acknowledging the achievement of standard search theory, we pointed out fundamental problems with the theory. Most serious is the assumption that the job arrival rate, the job separation rate, and the probability distribution of wages (more generally, some measure of the desirability of the jobs) are common to all the workers and firms. This assumption, though it is taken for granted in most models, is simply untenable. There is always the economy-wide distribution of economic variables of interest such as the job arrival rate, the job separation rate, and wages, of course. However, it is not relevant distribution for individual worker and firm. Each economic agent faces different job arrival rate, job separation rate and probability distribution of wages. In short, each economic agent acts in its own "universe". For this reason, even the unique reservation wage as a determinant of unemployment cannot be well defined. It is, in fact, frictions and uncertainty emphasized by the equilibrium search theory that makes the economy-wide distribution or the average irrelevant to economic decisions made by individual economic agent. The standard approach to take care of economic agents' heterogeneity by way of a single probability distribution common to all the agents is on the wrong track.

The concept of stochastic macro-equilibrium is motivated by the presence of all kinds of unspecifiable micro shocks. At first, one might think that allowing all kinds of unspecifiable micro shocks leaves so many degrees of freedom that almost anything can happen. However, the methods of statistical physics—the maximization of entropy under macro-constraints—actually provide us with the quantitative prediction about the equilibrium distribution of productivity, namely Eq. (27).

It is extremely important to recognize that the present approach does not regard behaviors of workers and firms as random. They certainly maximize their objective functions perhaps dynamically in their respective stochastic environments. The maximization of entropy under the aggregate demand constraint (6), in fact, balances two forces. On one hand, whenever possible, workers are assumed to move to better jobs which are identified with job sites with higher productivity. Firms make efforts for hiring good workers under demand constraint in the goods market. It is the outcome of successful job matching resulting from the worker's search and the firm's recruitment. When the level of aggregate demand is high, this force dominates because demand for labor of high productivity firms is high. However, as the aggregate demand gets lower, the number of possible allocations consistent with the level of aggregate demand increases. More workers are forced to be satisfied with or look for low productivity jobs. Randomness which plays a crucial role in our analysis basically comes from the fact that demand constraints in the product market facing firms with different productivity, and optimizing behaviors of workers and firms under such constraints are so complex and unspecifiable that those of us who analyze the macroeconomy must take micro behaviors as random. The method is straight-forward, and does not require any arbitrary assumptions on the behavior of economic agents.

When the level of aggregate demand is high, it is most likely that high productivity firms keep more workers on the job,[11] and make more aggressive hiring efforts than in the period of low demand. Workers are certainly aware of such a change. It is well known that quit rates are higher in high-demand periods despite of the fact that the employed workers are treated better in such periods.

We emphasize that *frictions and uncertainty in the labor market are not exogenously given, but depend crucially on the aggregate demand.* The aggregate demand, therefore, fundamentally conditions the rate of successful matching. The entropy maximization plays the role of matching function in standard search theory. We must note that the role of the reservation job attractiveness, μ in the model, also depends crucially on the aggregate demand. It is shown that in deep recession, the reservation job attractiveness (or wages) plays only a relatively small role as a determinant of unemployment.

Keynes' theory has been long debated in terms of unemployment or "involuntary" unemployment. Though unemployment is one of the most important economic problems in any society, to focus only on unemployment is inadequate for the purpose of providing micro-foundations for the Keynesian economics. The real issue is whether or not there is any room for mobilizing labor to high productivity jobs, firms, or sec-

---

[11] To be precise, more workers are on the "job sites" with higher productivity because productivity may differ across job sites even within a single firm. The method I used in this paper should apply to "job sites" rather than firms, because in general, a firm is a cluster of job sites with different productivity levels.

tors. The famous Okun's law demonstrates that there is always such a room in the economy (Okun 1963); See Syverson (2011) on more recent research on productivity dispersion.

Based on the methods of statistical physics, the present paper quantitatively shows how labor is mobilized when the aggregate demand rises. The level of aggregate demand is the ultimate factor conditioning the outcome of random matching of workers and monopolistically competitive firms. By so doing, it changes not only unemployment but also the distribution of productivity, and as a consequence, the level of aggregate output. This is the market mechanism beneath Keynes' principle of effective demand. Contrary to many economists' belief, the old principle of effective demand has solid micro-foundations. The market mechanism beneath Keynes' principle of effective demand is general equilibrium of monopolistic competition coupled with search by workers and firms under friction and uncertainty. The Keynesian economics, in effect, claims that in the short-run, aggregate demand ultimately conditions the matching of workers and firms, thereby determines the utilization of labor and the level of output in the macroeconomy. This logic does not depend on details of economic agents' micro behavior.

References

Aoki M, Yoshikawa H (2007) Reconstructing macroeconomics: a perspective from statistical physics and combinatorial stochastic processes. Cambridge University Press, Cambridge

Aoyama H, Yoshikawa H, Iyetomi H, Fujiwara Y (2010) Productivity dispersion: facts, theory, and implication. J Econ Interac Coord 5:27–54

Caballero RJ (2010) Macroeconomics after the Crisis: time to deal with the pretense-of-knowledge syndrome. J Econ Persp 24(4):85–102

Delli Gatti D, Gaffeo E, Gallegati M, Giulioni G, Palestrini A (2008) Emergent macroeconomics. Springer, Milan, Italy

Diamond P (2011) Unemployment vacancies, wages. Am Econ Rev 101(4):1045–1072

Drèze JH (1979) Demand estimation risk-aversion and sticky prices. Econ Lett 4:1–6

Feynman RP (1964) The principle of least action. Chapter 19 of his lectures on physics, vol II. Addison Wesley, Reading

Foley D (1994) A statistical equilibrium theory of markets. J Econ Theory 62:321–345

Friedman M (1968) The role of monetary policy. Am Econ Rev 58:1–17

Galibaldi U, Scalas E (2010) Finitary probabilistic methods in econophysics. Cambridge University Press, Cambridge

Ikeda Y, Souma W (2009) International comparison of labor productivity distribution for manufacturing and non-manufacturing firms. Prog Theor Phys Suppl 179:93–102

Iyetomi H (2012) Labor Productivity distribution with negative temperature. Prog Theor Phys Suppl 194:135–143

Iyetomi H, Nakayama Y, Yoshikawa H, Aoyama H, Fujiwara Y, Ikeda Y, Souma W (2011) What causes business cycles? Analysis of the Japanese iindustrial production data. J Jpn Int Econ 25:246–272

Keynes JM (1936) General theory of employment, interest, and money. Macmillan, London

Kittel C, Kroemer H (1980) Thermal physics, 2nd edn. W H Freeman, San Francisco

Kydland F, Prescott E (1982) Time to build and aggregate fluctuation. Econometrica 50:1345–1370

Marsil M, Zhang YC (1998) Interacting individuals leading to Zipf's law. Phys Rev Lett 80(12):2741–2744

Mortensen DT (2011) Markets with search friction and the DMP model. Am Econ Rev 101(4):1073–1091

Negishi T (1960-61) Monopolistic competition and general equilibrium. Rev Econ Stud 28:196–201

Negishi T (1979) Microeconomic foundation of Keynesian macroeconomics. North-Holland, Amsterdam

Okun A (1963) Potential GNP: its measurement and significance. In: Proceedings of the business and economic statistics section, American statistical association, pp 98–104. Reprinted in Okun's The Political Economy of Prosperity, Washington D.C.: The Brookings Institution, 1970

Okun A (1973) Upward mobility in a high-pressure economy. Brook Pap Econ Act 1:207–261

Pissarides CA (2011) Equilibrium in the labor market with search frictions. Am Econ Rev 101(4):1092–1105

Postel-Vinay Fabien, Robin Jean-Marc (2002) Equilibrium wage dispersion with worker and employer heterogeneity. Econometrica 70(6):2295–2350

Rogerson Richard, Shimer Robert, Wright Randall (2005) Search theoretic models of the labor market: a survey. J Econ Lit XLIII:959–988

Solow RM (1986) Unemployment: getting the questions right. Economica 53(Supplement):S23–34

Sweezy PM (1939) Demand under condition of oligopoly. J Polit Econ 47:568–573

Syverson C (2011) What determines productivity? J Econ Lit 49(2):326–365

Tobin J (1972) Inflation and unemployment. Am Econ Rev 62:1–18

Tobin J (1993) Price flexibility and output stability: an old Keynesian view. J Econ Persp 7:45–65

Yoshikawa H (2003) The role of demand in macroeconomics. Jpn Econ Rev 54(1):1–27

Yoshikawa H (2013) Stochastic macro-equilibrium: a microfoundation for the Keynesian economics. CIRJE-F-827. University of Tokyo, Faculty of Economics