American Journal of Information Systems, 2014, Vol. 2, No. 1, 6-10
Available online at http://pubs.sciepub.com/ajis/2/1/2

DOI:10.12691/ajis-2-1-2

Decision Making Problem in Division of Cognitive Labor with Parameter Inaccuracy: Case Studies

Jin Huan Zhang, Khin War War Htike, Ammar Oad, Hao Zhang*

School of Information Science and Engineering, Central South University, Changsha, Hunan, P.R. China
*Corresponding author: hao@csu.edu.cn

Received December 11, 2013; Revised December 17, 2013; Accepted December 29, 2013

Abstract Scientific communities will be more effective for society if scientists effectively divide their cognitive labor. So one way to study how scientists divide their cognitive labor has become an important area of research in science. This problem was firstly discovered and studied by Kitcher. Later on, Kleinberg and Oren pointed out that the model proposed by Kitcher might not be realistic. We investigate the impact of the imprecise parameter in project selection results. In this paper, we further our study on this issue. We study the policy of decision making problem based on the modified division of cognitive labor model with the assumption that a scientist is aware of the existence of the imprecise parameters and provide the detailed analytical results. And we provide a decision rule to minimize the possible loss based on error probability estimation.

Keywords: *cognitive labor, imprecise parameters*

Cite This Article: Jin Huan Zhang, Khin War War Htike, Ammar Oad, and Hao Zhang, "Decision Making Problem in Division of Cognitive Labor with Parameter Inaccuracy: Case Studies." *American Journal of Information Systems* 2, no. 1 (2014): 6-10. doi: 10.12691/ajis-2-1-2.

1. Introduction

The best known approach for modeling congnitive labor has been created by Philip Kitcher [1,2] and Michael Strvens [3,4]. Kitcher proposed that the progress of science will be optimized when there is an optimal distribution of cognitive labor within the scientific community. However, the main argument of the model proposed by Kitcher has made many assumptions that might not be realistic.

The division of cognitive labor model is a procedure by which scientists calculate their expected rewards. Scientists calculate their marginal contribution to the probability of the success of project and then uses this information to estimate the expected reward and payoff. They assumed that scientists are utility-maximizes, the division of cognitive labor among a number of pre-defined projects, (distribution assumption) every scientist knows the distribution of cognitive labor before they choose what project to work on and finally (success function assumption) each project has a success function, and takes the units of cognitive labor as input and the mainly objective is probabilities of success of project.

Kitcher found that, when scientists made their decisions out of their personal interests for awards, the result might be even better than the pure one. Kitcher employed a mathematical model to support his argument. We got a lot of lessons from past experiences because the seemed mostly unlikely projects (or theory) might be proved to be the correct one in the end. Kitcher mentions in his work, that when those high-minded goals are replaced by baser motives such as thirst for fame, some scientists will automatically choose the second project towards the improvement of the total probability of success.

The main argument of division of cognitive labor model is that the assumptions of the model are not realistic. In division of cognitive labor model, Kitcher and Strevens used a representative-agent approach that means they assumed that every individual agent who is exactly the same situation as all the others. In this research, we keep only the representative agent approaches like Kitcher because we emphasize only about success function assumption with parameter inaccuracy. Like Gilbert said, we overrate or underrate the odds that it will occur. And hence we overrate or underrate the actual value of the gain [5]. He gave a lot of examples to prove that there are errors at the odds and value of gain when we decide what the right thing is to do.

In this paper, we introduce perturbation into the basic model parameters and studied the policy decisions with its impact on the original distributions. The theoretical and experimental results demonstrate that, under the provided conditions, the distributions are different from those obtained in the ideal case.

The rest of the paper is organized as follows. In section II, the related works are introduced. We study analytically the policy decisions under the division model with imprecise parameters and derive the close-form conditions in section III. Section IV bestows some experimental results for a number of cases. Finally, section V concludes the paper.

2. Related Works

In this section, we presents others work which is related to this research. We examine and review some of the existing solutions on different aspects of division of cognitive labor model in science. The idea of division of cognitive labor model, due to Kitcher, is one of the most striking features of modern social community.

However, in recent time there are many researches that are pointing out the weakness of Kitchers model. There are a large number of approaches for learning the structure and community of social network. Nowadays, most of young scientists and agents are facing with decision problem. They choose a method or make a decision according to their own past experience, and also following their peer or neighbors. Bala and Goyal presented a very general model in economic [6]. Langhe and Grieff studied the division of cognitive labor in science [7]. They generalized that Kitchers conclusion about the division of cognitive labor in science is not robust against changes in his single standard view to multiple standards.

In recent times, Jon Kleinberg and Sigal Oren proposed an idea to improve the social optimum in scientific credit allocation [8]. They adapted Kitcher model and showed that the misallocating scientific credit mechanisms might be a good way to obtain the social optimality. However, at [8] they built the credit allocation model to choose one among projects of varying levels of importance and difficulty of projects.

The division of cognitive labor is one of the most conspicuous model in modern social community. There are a large number of approaches for learning the structure and community of social network. The division of cognitive labor model has also been argued by Weisberg and Muldoon's epistemic landscape approach [9] and Zollman's epistemic networks approach [10,11]. In [9], they considered the original model in [1] is too ideal, e.g., all the agents know the distribution of cognitive labor at all times and division of cognitive labor model is not robust to change in the distribution assumption. They built robustness model, where agents actually calculate their marginal contributions to their project success using a specific function.

In recent time, most of young scientists and agents are facing with decision problem. They choose a method or make a decision according to their own past experience, and also following their peer or neighbors. Social network community have been studied in a number of domains; in economic, Bala and Goyal presented a very general model [12] and also another class of approaches to simple model of herd behavior, see work by [13,14]. They also study social network community by following the basic model of DeGroot [15] with the setting of individual agents are connected in social network and update their beliefs repeatedly taking into account the averaged weight of their neighbors' opinions can arrive a shared opinion [16,17,18].

We reviews some of the existing solutions and on different aspects of analyzing theories developed by prior studies in the area of social science, hhowever, they do not provide any detailed theoretical analysis or remedy on the impact of the variation on model parameters [19]. We recently modeled the impact of imprecise parameters theoretically and we provide detailed analytical results for this issue [20].

3. Theoretical Analysis

The fundamental equation of division of cognitive labor model built by Kitcher is as follows [1]:

$$p_i(n) = \rho_i(1-e^{-kn}) \tag{1}$$

In this equation, ρ_i and k are all parameters and $k \geq 0$ is called the responsiveness. And assume that there are N scientists (denoted by $S_i, i = 0,1,...,N$) working on M projects (denoted by $PJ_i, i = 0,1,...,M$. $p_i(n)$ represents the probability of success when n scientists are working on PJ_i .

Each scientist is assumed to choose a project that maximizes his/her probability of success. When a project is successful, all scientists working on it would equally share the credit. From the aspect of each scientist, the principle of choice is assumed to be based on the reward that he/she might receive, i.e., a scientist would choose a project with the largest reward $p_i(n_i)/n_i$.

Kitcher provided analytical results in [1] for the basic model. However, in reality, it is not possible that model parameters are known precisely by each scientist. In this paper, we only consider the case when the parameter ρ_i deviates from its true value.

Although we conducted theoretical analysis on the impact of imprecise parameters [13], we do not give the right decision that a scientist should do in this situation. In this paper, we would focus on the decision making issue based on analytical models for some special cases.

3.1. One Scientist and Two Projects Case

For this case, the scientist is denoted by S1 and two projects are denoted by PJ1 and PJ2. The general model in Eq.1 can be rewritten as follows:

$$p_1(1) = \rho_1(1-e^{-k})$$

$$p_2(1) = \rho_2(1-e^{-k})$$

In real applications, the estimated model parameters $\hat{\rho}_1$, $\hat{\rho}_2$ might not be the true values. If we assume the measured ρ_2 contains some perturbation, i.e., $\hat{\rho}_1 = \rho_1$, $\hat{\rho}_2 = \rho_2 + x$ where x is the added perturbation, and the above equations can be rewritten as [13]

$$\hat{p}_1(1) = \rho_1(1-e^{-k})$$

$$\hat{p}_2(1) = (\rho_2 + x)(1-e^{-k})$$

Naturally, the perturbation x would have great impact on this model. When $\hat{p}_2(1) > \hat{p}_1(1)$, S1 will switch from PJ1 to PJ2 due to better estimated success probability. We use P_e to represent this probability of error, which can be denoted by the following equation:

$$\begin{aligned}P_e &= P(PJ1)\cdot P(PJ2 \mid PJ1) + P(PJ2)\cdot P(PJ1 \mid PJ2)\\ &= P(PJ1)\cdot P(\hat{p}_2(1) > \hat{p}_1(1) \mid PJ1)\\ &\quad + P(PJ2)\cdot P(\hat{p}_1(1) > \hat{p}_2(1) \mid PJ2)\end{aligned}$$

It can be rewritten as:

$$\begin{aligned}P_e &= P(PJ1)\cdot P(x > \rho_1 - \rho_2 \mid PJ1)\\ &\quad + P(PJ2)\cdot P(x < \rho_1 - \rho_2 \mid PJ2)\end{aligned}$$

It can be further rewritten as:

$$P_e = P(PJ1)\cdot\int_{\rho_1-\rho_2}^{+\infty} f_1(x)dx + P(PJ2)\cdot\int_{-\infty}^{\rho_1-\rho_2} f_2(x)dx$$

It can be seen that P_e is a function of $\rho_1-\rho_2$. In order to make the minimum probability of miscarriage justice, the above formula does derivation on $\rho_1-\rho_2$ as:

$$\frac{\partial \mathrm{P}_e}{\partial(\rho_1-\rho_2)} = -P(PJ1)f_1(\rho_1-\rho_2) + P(PJ2)f_2(\rho_1-\rho_2) .$$

So that the value of the derivative to zero, we obtain:

$$\frac{P(PJ1)}{P(PJ2)} = \frac{f_2(\rho_1-\rho_2)}{f_1(\rho_1-\rho_2)}$$

Without generality, we assume $\rho_1 > \rho_2$ and $f_1(x) = f_2(x)$. The scientist would choose a project with much larger reward. If the scientist would choose project PJ 1, the probability of error is written as:

$$P_{e1} = P(\hat{p}_2(1) > \hat{p}_1(1) = P(x > \rho_1-\rho_2) = \int_{\rho_1-\rho_2}^{+\infty} f(x)dx$$

Otherwise, when the scientist would choose project 2, the probability of error is:

$$P_{e2} = P(\hat{p}_1(1) > \hat{p}_2(1)) = P(x < \rho_1-\rho_2) = \int_{-\infty}^{\rho_1-\rho_2} f(x)dx$$

If random variable x follows the uniform distribution, i.e.,

$$f(x) = \frac{1}{b-a}(a \le x \le b)$$

We obtain:

$$P_{e1} = \int_{\rho_1-\rho_2}^{+\infty} f(x)dx = \int_{\rho_1-\rho_2}^{+\infty} \frac{1}{b-a}dx$$

$$P_{e2} = \int_{-\infty}^{\rho_1-\rho_2} f(x)dx = \int_{-\infty}^{\rho_1-\rho_2} \frac{1}{b-a}dx$$

(1) If ρ_1-$\rho_2 \le (a+b)/2$, $P_{e1} \ge P_{e2}$ would hold. The scientist would choose project 2 due to better estimated success probability.

(2) If ρ_1-$\rho_2 > (a+b)/2$, $P_{e1} < P_{e2}$ would hold. The scientist would choose project 1 to get much more reward or success probability.

If random variable x follows the normal distribution, i.e.,

$$f(x) = \frac{1}{\sqrt{2\pi\delta}} e^{(x-\mu)^2/(2\delta^2)}$$

We obtain:

$$P_{e1} = \int_{\rho_1-\rho_2}^{+\infty} f(x)dx = \int_{\rho_1-\rho_2}^{+\infty} \frac{1}{\sqrt{2\pi\delta}} e^{(x-\mu)^2/(2\delta^2)}dx$$

$$P_{e2} = \int_{-\infty}^{\rho_1-\rho_2} f(x)dx = \int_{-\infty}^{\rho_1-\rho_2} \frac{1}{\sqrt{2\pi\delta}} e^{(x-\mu)^2/(2\delta^2)}dx$$

(1) If ρ_1-$\rho_2 \le \mu$, we have $P_{e1} \ge P_{e2}$. In this case, the scientist would choose project 2 to work due to better estimated success probability.

(2) If ρ_1-$\rho_2 > \mu$, $P_{e1} < P_{e2}$ would hold. The scientist would choose project 1 to get much more reward or success probability.

3.2. Two Scientists and Two Projects Case

For this case, two scientists are denoted by S1 and S2; two projects are denoted by PJ1 and PJ2, respectively. The distribution between scientists and projects could be <2, 0>, <0, 2> and <1, 1> three cases. Without parameter perturbations, S1 and S2 would not both choose PJ2, i.e., the stable distribution cannot be < 0, 2 > for the case $\rho_1 > \rho_2$.

Next, we would do analysis from the two aspects of personal and social with parameter perturbations. Without generality, we assume the measured ρ_2 contains some perturbation, i.e., $\hat{\rho}_1 = \rho_1$, $\hat{\rho}_2 = \rho_2 + x$ where x is the added perturbation.

(1) If the distribution is <2, 0>, the success probability or reward of scientists S1 and S2 working on project PJ1 is written respectively as:

$$p_1(2) = \rho_1(1-e^{-2k})/2$$

The aggregate probability of success is:

$$p_a = \rho_1(1-e^{-2k})$$

For the case with parameter perturbations, the estimated success probability is as follows:

$$\hat{p}_1(2) = \rho_1(1-e^{-2k})/2$$

$$\hat{p}_2(2) = (\rho_2 + x)(1-e^{-0k})/2 = 0$$

The aggregate probability of success is:

$$p_a = \rho_1(1-e^{-2k})$$

Obviously, the parameter perturbations would not have impact on estimated success probability and the aggregate probability of success.

(2) If the distribution <0,2> holds, the success probability or reward of scientists S1 and S2 working on project PJ2 is written respectively as:

$$p_2(2) = \rho_2(1-e^{-2k})/2$$

The aggregate probability of success is:

$$p_a = \rho_2(1-e^{-2k})$$

The estimated success probability due to parameter perturbations is written as:

$$\hat{p}_2(2) = (\rho_2 + x)(1-e^{-2k})/2$$

The aggregate probability of success is:

$$p_a = (\rho_2 + x)(1-e^{-2k})$$

(3) If the distribution is <1,1>, the success probabilities of scientist S1 working on project PJ1 and scientist S2 working on project PJ2 are written respectively as:

$$p_1(1) = \rho_1(1-e^{-k})$$

$$p_2(1) = \rho_2(1-e^{-k})$$

The aggregate probability of success is:

$$p_a = \rho_1(1-e^{-k}) + \rho_2(1-e^{-k})$$

With parameter perturbations, the corresponding estimated probability of success are as follows:

$$\hat{p}_1(1) = \rho_1(1-e^{-k})$$

$$\hat{p}_2(1) = (\rho_2 + x)(1-e^{-k})$$

The aggregate probability of success is:

$$p_a = \rho_1(1-e^{-k}) + (\rho_2 + x)(1-e^{-k})$$

In general, we assume $\rho_1 > \rho_2$ and the scientist S1 would choose one project to work firstly. If the scientist S1 chooses the project PJ1, the scientist S2 would choose the project PJ1 or the project PJ2. The error probability of the scientist S2 choosing the project PJ1 or the project PJ2 is computed respectively by:

$$P_{e1} = P(\hat{p}_2(1) > \hat{p}_1(2)) = P(x > \rho_1(1+e^{-k})/2 - \rho_2)$$
$$= \int_{\rho_1(1+e^{-k})/2-\rho_2}^{+\infty} f(x)dx$$

$$P_{e2} = P(\hat{p}_1(2) > \hat{p}_2(1)) = P(x < \rho_1\left(1+e^{-k}\right)/2 - \rho_2)$$
$$= \int_{-\infty}^{\rho_1\left(1+e^{-k}\right)/2-\rho_2} f(x)dx$$

If random variable x follows the uniform distribution, i.e.,

$$f(x) = \frac{1}{b-a}(a \le x \le b)$$

We obtain:

(1) If $\rho_1(1+e^{-k})/2 - \rho_2 \ge (a+b)/2$, we obtain $P_{e1} \le P_{e2}$. For this case, the scientist S2 would choose project PJ1 to work due to better success probability or reward. And the aggregate probability of success is:

$$p_{a1} = \rho_1(1-e^{-2k})$$

(2) If $\rho_1(1+e^{-k})/2 - \rho_2 < (a+b)/2$, $P_{e1} > P_{e2}$ would hold. The scientist S2 would choose project PJ2 to work on. The aggregate probability of success is:

$$p_{a2} = \rho_1(1-e^{-k}) + (\rho_2 + x)(1-e^{-k})$$

Next, it is also meaningful to study from the social aspect, i.e., if the scientist S2 would choose one of project to work on due to his/her better personal reward or success probability, is it the CO-distribution? For the distribution <2, 0>, only if $p_{a1} \ge p_{a2}$, it is the CO-distribution. When $p_{a1} \ge p_{a2}$, $x \le \rho_1 e^{-k} - \rho_2$ should hold. And the random variable x follows the uniform distribution, i.e., $a \le x \le b$.

Hence, the inequality $a < \rho_1 e^{-k} - \rho_2 \le b$ should hold. Considering the inequality $\rho_1(1+e^{-k})/2 - \rho_2 \ge (a+b)/2$, the scientist S2 would make the choice for the benefit of both personal and social interests.

If random variable x follows the normal distribution, i.e.,

$$f(x) = \frac{1}{\sqrt{2\pi\delta}} e^{(x-\mu)^2/(2\delta^2)}$$

Similarly, we obtain:

If $\rho_1(1+e^{-k})/2 - \rho_2 \ge \mu$, the inequality $P_{e1} \le P_{e2}$ would hold. The scientist S2 would choose project PJ1 due to better personal reward.

If $\rho_1(1+e^{-k})/2 - \rho_2 < \mu$, $P_{e1} > P_{e2}$ holds. S2 will make a choice on PJ2. If S2 would choose PJ1 due to better success probability, while making the distribution <2,0> be the co-distribution, the inequality $x \le \rho_1 e^{-k} - \rho_2$ should hold.

From the above analysis, we know that when the perturbation item x would meet certain conditions both the personal and social better success probability could be satisfied.

3.3. One Scientist and M Projects Case

Without generality, let we assume the probabilities of success for M projects satisfy $p_1 > p_2 > p_3 > ... > p_M$. S1 will choose PJ1 with the highest probability of success. If $p_i(i>1)$ contains perturbation item x, S1 will make the choice according to the following computation.

$$P_{ej} = P(PJj \mid PJ1) = P(P(PJj) > P(PJ1)) \quad j = 1, 2, 3, ..., M$$

The scientist S1 would choose project j only if P_{ej} is much smaller than others.

4. Simulation

We study the policy decision based on the modified division of cognitive labor model with the detailed simulation results. We would like to provide the probability of error when the scientist S2 chooses the projects with the assumption of parameter inaccuracy at the projects' probability of success. In this simulation experiment, the parameter values are as follows: $\rho_1 = 0.9, \rho_2 = 0.5, k = 0.4$ and the parameter inaccuracy x is set in random value between 0 and 1.

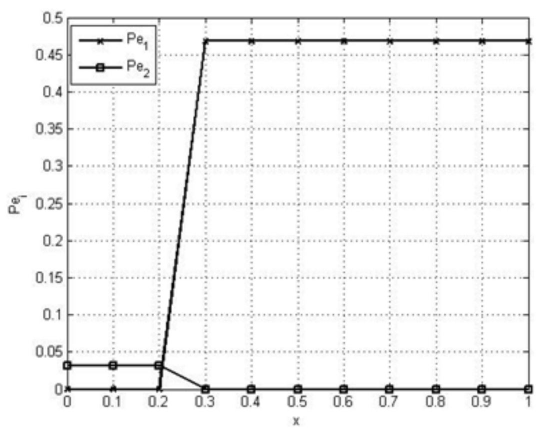

Figure 1. The error probability of the scientist S2 choosing the projects if *x* follows the uniform distribution

Figure 1 shows the probability of error according to the scientist S2 choice. When the scientist S1 is working in project PJ1, there are 2 ways to choose for scientist S2; he would follow <1,1> or <2,0> distribution. According to the result showing in Fig. 1, there would be only too small error probability if scientist2 chooses the project PJ2.

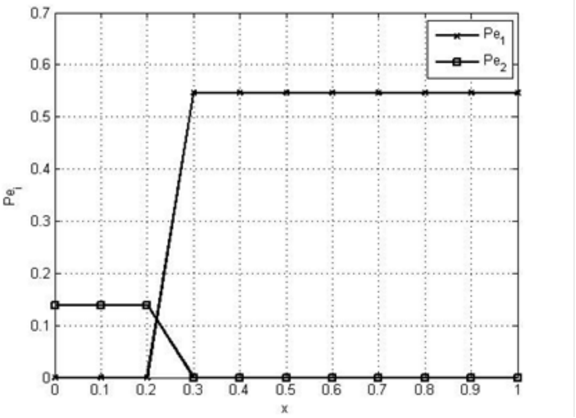

Figure 2. The error probability of the scientist S2 choosing the projects if x follows the standard normal distribution

Figure 2 shows that the error probability for scientist S2 with the assumption of parameter inaccuracy *x* follows the standard normal distribution. If scientist S2 chooses the project PJ1 to work together with scientist S1, he/she could get the probability of error over 0.5.

5. Conclusions

In this paper, we offer decision making rules in the division of cognitive labor with theoretical analysis on the impact of parameter perturbations. Due to the complexity involved in modeling the general case, we mainly focus on several special cases with a small number of scientists and projects. The analytical results demonstrate that when we assume the existence of parameters inaccuracy, we would choose the project that minimizes the error probability.

Acknowledgement

The work is supported by National Natural Science Foundation of China (No.61003033) and Specialized Research Fund for the Doctoral Program of Higher Education (No.20100162120018), the Hunan soft science project (2010ZK3061), Central Universities Fundamental Research Funds for the project (2012QNZT063).

References

[1] P. Kitcher, “The division of cognitive labor, journal of philosophy,” *The Journal of Philosophy*, vol. 87, no. 1, pp. 5-22, 1990.

[2] P. Kitcher, “The advancement of science,” *New York: Oxford University Press*, 1993.

[3] M. Strevens, “The role of the priority rule in science”, *Journal of Philosophy*, vol. 100, no. 2, pp. 55-79, 2003.

[4] M. Strevens, “The role of the matthew effect in science,” *Studies in the History and Philosophy of Science*, vol. 37, pp. 159-170, 2006.

[5] D. Gilbert, “Buried by bad decisions,” *Nature*, pp. 275-277, 2011.

[6] V. Bala and S. Goyal, “Learning from neighbours,” *Review of Economic Studies*, pp. 595-621, 1998.

[7] DL.Rogier and G.Matthias, “Standards and distribution of cognitive labor,” *Logic Journal of the IGPL*, 2009.

[8] J.Kleinberg and S.Oren, “Mechanisms for (mis)allocating scientific credit,” *Proceedings of the 43rd annual ACM symposium on Theory of Computing*, 2011.

[9] M. Weisberg and R. Muldoon, “Epistemic landscapes and the division of cognitive labor,” *Philosophy of Science*, pp. 37:159-170, 2009.

[10] K. J. Zollman, “The communication structure of epistemic communities,” *Philosophy of Science*, vol. 74(5), pp. 574-587, 2007.

[11] K. J. Zollman, “The epistemic benefit of transient diveristy,” *Erkenn*, vol. 72, pp. 17-35, 2010.

[12] Bala, Venkatesh and Sanjeev Goyal. "Learning from neighbours." Review of Economic Studies, pp. 595-621, 1988.

[13] Banerjee, A. V. "A simple model of herd behavior." The Quarterly Journal of Economics, 107 (3), pp. 797-817, 1992.

[14] Welch, I. "Sequential sales, learning, and cascades." The Journal of Finance, 47(2), pp. 695-732, 1992.

[15] M. H. DeGroot. "Reaching a consensus." J. American Statistical Association, 69, 118-121, 1974.

[16] B. Golub and M.O. Jackson. "Naive learning in social networks: Convergence, influence and the wisdon of crowds." American Econ. J.: Microeconomics, 2000.

[17] P. M. DeMarzo, D. Vayanos and J. Zweibel. "Persuasion bias, social influence, and unidimensional opinions." Auarterly Journal of Economics, 118(3), 2003.

[18] M. O. Jackson. "Social and Economic Networks." Princeton University Press, 2008.

[19] R. Muldoon and M.Weisberg, “Robustness and idealization in models of cognitive labor synthese,” Synthese, vol. 183, pp. 161-174, 2011.

[20] K. W. W. Htike and H. Zhang, “Theoretical study of division of cognitive labor with imprecise model parameters,” CSSS, pp. 4094-4097, 2012.