Model of endogenous welfare stigma: Statistical discrimination view

Kurita, Kenichi and Hori, Nobuaki and Katafuchi, Yuya

Kyushu University, Kyushu University, Research Institute for Humanity and Nature

1 February 2019

Stigma model of welfare fraud and non-take-up: Theory and evidence from OECD panel data

Kenichi Kurita[†]

Nobuaki Hori[‡]

Yuya Katafuchi[§]

Abstract

This paper tries to challenge two puzzles in the welfare benefit program. The first puzzle is 'non-take-up welfare' which means poor people do not take-up welfare even though they are approved to take-up. Second, empirical evidence suggests that there may exist the inverse U-shaped relationship between benefit level and beneficiary ratio. We present a model of welfare stigma as a hypothesis to explain the above puzzles. Specifically, we investigate the statistical discrimination view model. Results are summarized as the relationship between two types of elasticity.

Keywords: Stigma, Take-up, Minimum income guarantee, OECD panel data, Poverty

JEL classification: H31, H53, I38

[†]Corresponding author. Faculty of Economics, Kyushu University
Address: 744 Motooka, Nishi-ku, Fukuoka, 819-0395, Japan
Email: kurita.econ@gmail.com
[‡]Faculty of Economics, Kyushu University
Address: 744 Motooka, Nishi-ku, Fukuoka, 819-0395, Japan
Email: hori@econ.kyushu-u.ac.jp
[§]Research Institute for Humanity and Nature
Address: 457-4 Motoyama, Kamigamo, Kita-ku, Kyoto 603-8047 Japan
Email: katafuchi@chikyu.ac.jp

1 Introduction

To analyze welfare programs, most public economics researchers exploit a labor supply model that is based on the maximization problem of leisure and consumption goods. This model can explain 'welfare fraud' but not 'non-take-up of welfare': welfare fraud means households take-up even though they are non-eligible; non-take-up welfare signifies that needy poor people do not take-up welfare even though they are approved to take-up. However, 'non-take-up' occurs in most developed countries (Currie, 2006; Immervoll, 2009; Plueger, 2009). Moreover, the result in comparative statics in the standard model, an increase in the level of welfare benefit has a disincentive effect on labour supply and the number of welfare benefit increases. Our empirical evidence, which is described later, nevertheless highlights the ratio of recipients to population decreases in benefit level when the level is sufficiently high.

That is to say; there exist some factors that are not considered in the standard model. One such factor is stigma, a sociological concept describing a negative label applied to behavior by society or a social group. In particular, stigma is an important concept in social psychology (Major et al., 2018).

Moffitt (1983) conducted one of the earlier studies to focus on welfare stigma in economics by analyzing household decision-making regarding whether to take up welfare benefits or supply labor by including the stigma as a kind of monetary cost. Moreover, that paper empirically examined theoretical results using panel study of income dynamics (PSID). Consequently, that author suggested that fixed stigma is statistically significant, but that variable stigma with respect to benefit level is not.

Besley and Coates (1992) pioneering research analyzed situations wherein stigmas were endogenized. They presented two models of social stigma: statistical discrimination and taxpayer resentment. Their results indicated the occurrence of welfare fraud. As needy types usually chose to take-up welfare benefits, non-take-up of welfare benefits did not manifest in their model. However, take-up rate in the United Kingdom was approximately 80 % (Duclos, 1995), approximately 60-67 % in the United States (Blank and Ruggles, 1996), approximately

37 % in Germany (Riphahn, 2001) and 16.3-19.7 % in Japan (Tachibanaki and Urakawa, 2006). Blumkin et al. (2015) analyzed welfare stigma as a policy tool, which was used to restrain welfare fraud. Thus, non-take-up welfare did not manifest in their model[1].

This study extends the model of Besley and Coate (1992) to explain the occurrence of non-take-up of welfare benefits. Unlike Besley and Coate (1992), we endogenize decision-making for needy poor people. Our comparative analysis indicates that an increase in the benefit level makes non-take-up of welfare benefits more serious.

The structure of this paper is as follows; the next section shows some empirical evidence regarding the relationship the between benefit level and the recipient ratio. The third section presents the model and the basic setting. The forth section conducts comparative static analysis. The final section concludes this paper.

2 Some Empirical Evidence

This section presents empirical evidence to explore the relationship between the recipient ratio and the minimum income benefit level using the OECD panel data.

2.1 Econometric Model

The panel data were analyzed to investigate the correlation between the minimum guaranteed income level and social benefit recipients. The decision to employ the panel data to investigate the relationship reflects three motivations. First, a panel data model can have better prediction accuracy than the cross-sectional model and time-series model because it has more observations than cross-section data and time-series data. Second, it enables researchers to address the issue of endogeneity caused by omitted variable bias. Third, it allows us to include changes in society in the empirical analysis (Greene, 2012). This paper analyzes the relationship between the minimum income benefit level and social benefit recipient ratio

[1] Hupkau and Maniquet (2018) analyzed the problem of non-take-up of welfare from the perspective of identity economics (Akerlof and Kranton, 2000; Kranton, 2016)

based on the baseline model:

$$y_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + e_{it}, \tag{1}$$

where y_{it} is the dependent variable, $\mathbf{x}'_{it}$ is the K-dimensional vector of predictors consisting of the target explanatory variable and the covariates, $\boldsymbol{\beta}$ is the K-dimensional vector of unknown parameters, and e_{it} is the disturbance term, which is distributed as $e_{it} \sim \mathcal{N}(0, \sigma_e^2)$. Furthermore, in equation (1), $i = 1, \ldots n$ indicates the index for a country, whereas $t = 1, \ldots, T$ represents the index for time. The OLS estimation of equation (1) after pooling the available data is called the pooling estimation.

When we consider the country-specific heterogeneity in the disturbance term of equation (1), e_{it} can be decomposed as follows:

$$\begin{aligned} y_{it} &= \mathbf{x}'_{it}\boldsymbol{\beta} + e_{it} \\ e_{it} &= \alpha_i + \nu_{it}, \end{aligned} \tag{2}$$

where α_i is the error depending on the country i and $\nu_{it} \sim$ i.i.d. $\mathcal{N}(0, \sigma_\nu^2)$ is the stochastic disturbance term. Equation (2) can be considered a one-way error component model (Baltagi, 1984) because it decomposes the disturbance term e_{it} into the error based on the individual heterogeneity and the stochastic error. The model in equation (2) can be estimated using a one-way fixed-effect estimator (hereinafter, one-way FE) and the one-way random-effect estimator (hereinafter, one-way RE). The one-way FE presumes the binary dummy variable for α_i whereas the one-way RE assumes that the individual effect is randomly determined.

Considering the heterogeneity caused by the individual effect as in equation (1), the disturbance term can be further decomposed to incorporate heterogeneity in time:

$$\begin{aligned} y_{it} &= \mathbf{x}'_{it}\boldsymbol{\beta} + e_{it} \\ e_{it} &= \alpha_i + \lambda_t + \nu_{it}, \end{aligned} \tag{3}$$

where λ_t is the error depending on the time t. Equation (3), a two-way error component model (Baltagi, 1984), decomposes the disturbance term into the error based on the heterogeneity of country i, the error caused by the time such as economic shocks, and the stochastic disturbance. As with equation (2),the model of equation (3) can be estimated by a two-way fixed-effect estimator (hereinafter, two-way FE) and a two-way random-effect estimator (hereinafter, two-way RE).

This paper estimates the relationship between the minimum income benefit level and social benefit recipients using five estimation methods: pooling, one-way FE, one-way RE, two-way FE, and two-way RE. These estimation methods are assessed via hypothesis testing. We first implement the F-test for pooling versus one-way FE or two-way FE. Second, we perform the Lagrange multiplier test (hereinafter, LM-test) (Honda, 1985) for pooling versus one-way RE or two-way RE. Finally, we conduct a Hausman test (Hausman, 1978) for one-way RE versus one-way FE, two-way RE, and two-way FE. Further information on hypothesis testing in the panel data analysis has been given by Baltagi (2008).

2.2 Data

This section proposes the detail of our dataset used for estimation of the panel data models introduced in Section 2.1. All of the data described below were obtained from OECD.Stat (OECD, 2019).

For the dependent variable, we use the logit-transformed version (`logit_recipients_ratio`) of the recipients ratio (`recipient_ratio`), which is the ratio of social benefit recipients to the total population. Data on the number of social benefit recipients were retrieved from the Social Benefit Recipients Database, and total population data were obtained from Population Statistics.

For the target explanatory variable, we include the minimum guaranteed income `mgincome`, which represents the degree of social benefits in terms of the ratio of the per capita social benefits to the median per capita income. These data can be retrieved from the Adequacy of Guaranteed Minimum Income Benefits. Furthermore, we incorporate the quadratic term `mgincome` (`mgincome_2`) to consider the nonlinear effect of the target explanatory variable.

In order to account for any estimation biases caused by unobserved confounders, we additionally incorporate the following covariates into the vector of predictors:

- `log_gdp_capita`: the natural logarithm of GDP per capita (`gdp_capita`), retrieved from Annual National Accounts.
- `youth_dependency`: ratio of young population (0 to 14 years old) to productive population (15 to 64), retrieved from Population Statistics.
- `old_dependency`: ratio of old population (over 65 years old) to productive population (15 to 64), retrieved from Population Statistics.
- `divorce_rate`: the marriage divorce rate, retrieved from Family Database.
- `unemployment`: the national unemployment rate for working-age population, retrieved from Labor Force Statistics.

The panel dataset using a date on the aforementioned variables. After reducing some missing series in the sample that was not randomly missing, we obtain panel data on $n = 25$ countries covering the time frame 2007 to 2012. This paper conducts the empirical analysis using the panel data with the number of observation $nT = N = 150$.

2.3 Result

This section presents the result of the empirical analysis investigating the relationship between the minimum guaranteed income level on the ratio of the number of recipients.

Table 1 presents the descriptive statistics of pooled panel data. This table demonstrates the large inequality between the minimum and maximum recipient ratio (minimum: 0.001, maximum: 0.037). Furthermore, the maximum of `mgincome` in Table 1 indicates that countries tend to guarantee almost 60% of the median per capita income through its social benefit programme, although the median and mean of the guaranteed minimum income is about 40%.

Examining the descriptive statistics by country, Table 2 indicates the necessity of adjustment by covariates or dealing with country-based heterogeneity when we assume that the minimum income benefit level is the determinant factor influencing benefit recipients/ total population ratio. For example, Canada and the Slovak Republic have the same maximum mean of recipient rate (0.034); however, their mean minimum guaranteed income level differs (Canada: 0.368, Slovak Republic: 0.238).

Table 3 presents the descriptive statistics by year. Although no large difference in means and medians can be found in this table, the standard deviation of the minimum guaranteed income level has a relatively large outlier in 2012 (0.89). This motivates us to include time-specific heterogeneity into our model by estimating the two-way error component model.

Before proceeding to regression analysis, let us discuss the simple correlation between benefit level and recipient ratio. Figure 1 presents the scatter plot of the observed couples (`mgincome`, `recipient_ratio`). Even though the figure depicts the roughly convex relationship of two variables in interest, possible confounders might lead to a spurious correlation among them. We thus discuss a regression analysis taking into account other factors, which may affect both of these target variables, and unobserved heterogeneity pertaining to country-specific factors and time-specific factors.

As the main findings in this empirical evidence, Table 4 shows the estimation results of panel data regression models based on the data introduced in Section 2.2. Each row corresponds to an explanatory variable, and each column corresponds to an estimation method. The standard errors of the estimated coefficients are estimated using the heteroskedasticity and autocorrelation consistent estimator (hereinafter HAC estimator) of Arellano (1987). The bottom part of this table gives the results of the hypothesis testing carried out for model evaluation.

Regarding the hypothesis testing concerning the pooling estimation, both one-way FE and two-way FE are accepted at 1% statistical significance according to the F-test results. LM-tests for the random-effect estimators reject the pooling estimation at 1% significance but accept the one-way RE and two-way RE at the same level of significance. In the comparison of fixed-effect estimators and random-effect estimators, Hausman tests do not reject either one-way RE or two-way RE. Furthermore, neither of the fixed-effect estimators are accepted.

Looking at the estimated coefficients by pooling estimation, `mgincome` has a significantly positive effect on the recipient ratio, and its quadratic term has a significantly negative effect on the recipient ratio. This suggests that the minimum guaranteed income level has an upper convex effect on recipient/population ratio. However, the results of F-test, which compares the pooling estimation with the fixed-effect estimators, and of the LM-test, which compares the pooling estimation with the random-effect estimators, highlight the necessity to take heterogeneity in a country or in both a country and time into account.

The Hausman test results in Table 4 suggest that the correlation between the explanatory variables and country effect or between the explanatory variables and both country effect and time effect is not statistically significant, i.e., the correlation between $\mathbf{x}_{it}$ and α_i or $\mathbf{x}_{it}$ and both α_i and λ_t is not statistically significant. Therefore, the random-effect estimator, which assumes no correlation between the explanatory variables and decomposed effects such as α_i and λ_t, is the most preferable method according to the hypothesis test results. In the estimation result of one-way RE considering country-specific heterogeneity, the minimum guaranteed income level has an upper convex effect on recipient/population ratio as well as the pooling estimation. This relationship is similar to the one found in the estimation of the two-way error component models.

Figure 2 presents the fitted curve of one-way RE with the scatter plot of the couples of observations (`mgincome`, `logit_recipient_ratio`). As we discussed, the one-way RE curve indeed visually indicates the upper convex relationship between the benefit level and the beneficiary ratio. Moreover, the figure gives two categories and members defined by an estimated maximum value of recipient ratio. For example, `group_2`, whose members have more `mgincome` than the benefit level corresponding to an estimated maximum of `logit_recipient_ratio`, includes Netherlands, Denmark and Germany.

The empirical results presented in this section have demonstrated the existence of an upper convex relation between the benefit level and the recipient ratio.

3 Model

In this section, we present a model of welfare stigma to explain 'non-take-up'.

3.1 The basic setting

There are two types in the economy. A 'needy type' is an individual who cannot work and a 'non-needy type' is defined as an individual who can work if he or she hopes so. We assume that a proportion of needy types in the total population is $\gamma \in (0,1)$. In the economy, needy types are eligible for welfare benefits, and non-needy types are not. That is, it is called 'non-take-up welfare' that the needy type does not take-up welfare benefit and 'welfare fraud' that the non-needy type take-up welfare benefit. To make the notation clear, we denote the needy type as 'type 1' and the non-needy as 'type 2'. Type 1 individuals have two choices; take-up welfare or not. The utility setting is,

$$\begin{cases} u\left(b, z_1\right) - \phi_1 s\left(p, q, z_1\right) & \text{if taking up welfare,} \\ 0 & \text{otherwise,} \end{cases}$$

where s is an index of stigma cost, which is explained later, p is a proportion of recipients to sub-population in type 1, q is a proportion of recipients to sub-population in type 2 and $\phi_i$is the sensitivity to stigma which varies over type i's sub-population, $\phi_i \tilde{U}[0,\phi]$, ϕ_i and ϕ_j are i.i.d, $i,j = 1,2, ij$. $u(\cdot,\cdot)$ denotes a material utility, z_i is type i's capability of consumption, $i = 1,2$. b is a level of welfare benefit. We assume the following properties, $\forall z_i$, $i = 1,2$, an income $I \in w, b$[2]

[2]For simplicity, we assume the price of consumption good is 1.

$$\frac{\partial u(I,z_i)}{\partial I} > 0,$$
$$\frac{\partial u(I,z_i)}{\partial z_i} > 0,$$
$$\frac{\partial u(I,z_i)}{\partial I \partial z_i} \geq 0.$$

The third property means that capability and consumption are complementary.

Type 2 individuals have two choices to either accept welfare benefits or work. Type 2's utility setting is as follows:

$$\begin{cases} u\left(b, z_2\right) - \phi_2 s\left(p, q, z_2\right) & \text{if taking up welfare,} \\ u\left(w, z_2\right) - \theta & \text{if working.} \end{cases}$$

Here θ is disutility of labor, and w is work income. We assume $z_2 > z_1$, that is to say, type 2's capability is higher than that of type 1 individuals cannot work because of time constraints, physical disabilities or mental illness. These constraints can affect consumption. For example, it makes sense that a single-parent household with limited free time will not enjoy consumption from income I less than a parent's household.

3.2 The critical level of sensitivity to stigma

To understand a household's decision-making, we consider the critical sensitivity of stigma cost, ϕ_i, as follows:

$$u\left(b, z_1\right) - \hat{\phi}_1 s\left(p, q, z_1\right) = 0,$$
$$u\left(b, z_2\right) - \hat{\phi}_2 s\left(p, q; z_2\right) = u\left(w, z_2\right) - \theta.$$

A type 1 household, where ϕ_1 is less than or equal to $\hat{\phi}_1$ prefers to take-up welfare. Then, all households in which $\phi_1 \in \left[0, \hat{\phi}_1\right]$ choose to take-up welfare and all households in which $\phi_1 \in \left(\hat{\phi}_1, \overline{\phi}\right]$ do not. Similarly, type 2 households in which ϕ_2 is less than or equal to $\hat{\phi}_2$ prefer to take-up welfare. All households in which $\phi_2 \in \left[0, \hat{\phi}_2\right]$ choose to take-up welfare On the other hand, all households in which $\phi_1 \in \left(\hat{\phi}_1, \overline{\phi}\right]$ choose to work.

The proportion of recipients in type 1, p, is as follows:

$$p = \min\left\{\frac{\hat{\phi}_1}{\overline{\phi}}, 1\right\} = \min\left\{\frac{u\left(b, z_1\right)}{\overline{\phi}s\left(p, q, z_1\right)}, 1\right\}.$$

And the proportion of recipients in type 2, q, is as follows:

$$q = \min\left\{\frac{\hat{\phi}_2}{\overline{\phi}}, 1\right\} = \min\left\{\frac{u\left(b, z_2\right) - u\left(w, z_2\right) + \theta}{\overline{\phi}s\left(p, q, z_2\right)}, 1\right\}.$$

3.3 Formulation of the stigma cost function

In this section, we formulate the stigma cost function. The probability that the recipients are non-needy is given by the following:

$$\Pr(i = 2|\text{Take-up welfare}) = \frac{(1-\gamma)q}{\gamma p + (1-\gamma)q} := \Pi.$$

We assume that stigma cost is an increasing function with π as follows:

$$\begin{aligned} &s = s\left(\Pi(p, q), z_i\right), \\ &\frac{\partial s\left(\Pi(p, q), z_i\right)}{\partial \Pi} > 0, \text{ for } i = 1, 2. \end{aligned}$$

This formulation is inspired by the statistical stigma in Besley and Coate (1992) and Blumkin et al. (2015). Setting a stigma means as follows. People in society despise 'welfare fraud' (the taking-up welfare by non-needy type (type 2)). However, without distinguishing between type 1 and 2, it is difficult to know whether welfare fraud is actually being committed.

Stigma cost is a function of capability. While Besley and Coate (1992) assumed that stigma cost was the same for all recipients, we differentiate stigma cost by the capabilities of type 1 and 2. Even though, we do not assume the sign of $\partial s\left(\Pi, z_i\right)/\partial z_i$. We denote π as the ratio p/q, then,

$$\Pi = \frac{1}{\gamma p/(1-\gamma)q + 1} = \frac{1}{\gamma/(1-\gamma)\pi + 1}.$$

I can rewrite this as follows:

$$s = s\left(\Pi(p,q), z_i\right) = s\left(\Pi(p/q,1), z_i\right) := s\left(\pi, z_i\right).$$

Clearly, we obtain the following:

$$\frac{\partial s\left(\Pi, z_i\right)}{\partial \Pi}\frac{\partial \Pi}{\partial \pi} < 0.$$

An equilibrium point corresponds to a solution in the following simultaneous equation:

$$\begin{cases} p = \dfrac{\hat{u}\left(b, z_1\right)}{\overline{\phi} s\left(\pi, z_1\right)}, \\ q = \dfrac{\hat{u}\left(b, z_2\right)}{\overline{\phi} s\left(\pi, z_2\right)}, \\ \pi = \dfrac{p}{q}. \end{cases}$$

Substituting the first and the second row equations into the right hand side of the third row equation, it indicates

$$\begin{aligned} \pi &= \frac{p(\pi)}{q(\pi)} \\ &= \frac{\hat{u}\left(b, z_1\right)}{\hat{u}\left(b, z_2\right)}\frac{s\left(\pi, z_2\right)}{s\left(\pi, z_1\right)} := M(\pi). \end{aligned}$$

Here,

$$\hat{u}\left(b, z_1\right) \equiv u\left(b, z_1\right),$$

$$\hat{u}\left(b, z_2\right) \equiv u\left(b, z_2\right) - u\left(w, z_2\right) + \theta.$$

$\hat{u}(b, z_i)$ is the incremental material utility when taking-up welfare. $M(\pi)$ is a mapping from

π to itself. By differentiation, we obtain the following:

$$\begin{aligned}
\frac{dM(\pi)}{d\pi} &= \frac{\hat{u}(b, z_1)}{\hat{u}(b, z_2)}\left[\frac{\partial s(\pi, z_2)/\partial\pi}{s(\pi, z_1)} - \frac{s(\pi, z_2)}{s(\pi, z_1)^2}\frac{\partial s(\pi, z_1)}{\partial\pi}\right] \\
&= \frac{\hat{u}(b, z_1)}{\hat{u}(b, z_2)}\frac{s(\pi, z_2)}{s(\pi, z_1)}\left[\frac{\partial s(\pi, z_2)/\partial\pi}{s(\pi, z_2)} - \frac{\partial s(\pi, z_1)/\partial\pi}{s(\pi, z_1)}\right] \\
&= \frac{\partial s(\pi, z_2)}{\partial\pi}\frac{\pi}{s(\pi, z_2)} - \frac{\partial s(\pi, z_1)}{\partial\pi}\frac{\pi}{s(\pi, z_1)}.
\end{aligned}$$

Here, we define the elasticity of stigma cost to π:

$$\varepsilon_\pi(z_i) \equiv -\frac{\partial s(\pi, z_i)}{\partial\pi}\frac{\pi}{s(\pi, z_i)}.$$

Using this elasticity, we rewrite this as given:

$$\frac{dM(\pi)}{d\pi} = \varepsilon_\pi(z_1) - \varepsilon_\pi(z_2).$$

Equation 3.3 corresponds to a slope of $M(\pi)$, which is a change of ratio to itself. Then, if $\varepsilon_\pi(z_1) - \varepsilon_\pi(z_2)$ in some domain, the possibility of multiple equilibria exists. The stability condition is

$$\varepsilon_\pi(z_1) - \varepsilon_\pi(z_2) < 1.$$

4 Comparative statics

In this section, we conduct comparative statics. We are particularly interested in how a change in benefit level to equilibrium and we compare our empirical evidence and theoretical results.

We define the elasticity as follows:

$$\eta_b(z_i) \equiv \frac{\partial\hat{u}(b, z_i)}{\partial b}\frac{b}{\hat{u}(b, z_i)}.$$

This is an elasticity of material utility to benefit level. The result is summarized in the following proposition.

Proposition 1

$$\begin{aligned}
\operatorname{sgn}\left[\frac{dp^*}{db}\right] &= \operatorname{sgn}\left[\frac{\eta_b(z_1)}{\eta_b(z_2)} - \frac{\varepsilon_{\pi^*}(z_1)}{1+\varepsilon_{\pi^*}(z_2)}\right], \\
\operatorname{sgn}\left[\frac{dq^*}{db}\right] &= \operatorname{sgn}\left[\frac{\eta_b(z_1)}{\eta_b(z_2)} - \frac{1+\varepsilon_{\pi^*}(z_1)}{\varepsilon_{\pi^*}(z_2)}\right], \\
\operatorname{sgn}\left[\frac{d\pi^*}{db}\right] &= \operatorname{sgn}\left[\frac{\eta_b(z_1)}{\eta_b(z_2)} - 1\right].
\end{aligned}$$

Proof. See appendix. ■

When the ratio, $\eta_b(z_1)/\eta_b(z_2)$, is sufficiently low, the equilibrium proportion of recipient in the needy type, p^*, the equilibrium proportion of recipient in the non-needy type, q^*, and the ratio of them, $\pi^* = p^*/q^*$, decrease in the level of welfare benefit.

In the second section, we show some empirical evidence that there exists an upper convex relation between the benefit level and the recipient ratio. Let denote R as a proportion of recipients to total population. Since the size of population is normalized to 1, R is given as follows:

$$R = \gamma p + (1-\gamma)q.$$

An effect of a change in benefit level on R is:

$$\frac{dR^*}{db} = \gamma\frac{dp^*}{db} + (1-\gamma)\frac{dq^*}{db}.$$

The sign of $\frac{dR^*}{db}$ is:

$$\operatorname{sgn}\frac{dR^*}{db} = \operatorname{sgn}\left[\frac{\eta_b(z_1)}{\eta_b(z_2)} - \frac{1-\gamma+\varepsilon_{\pi^*}(z_1)}{\gamma+\varepsilon_{\pi^*}(z_2)}\right].$$

The recipient ratio increases in the benefit level when the ratio of elasticity, $\eta_b(z_1)/\eta_b(z_2)$, is sufficiently low, vice versa. We present a model as a hypothesis to explain some empirical evidence; a raise of the benefit level reduces the recipient ratio when the benefit level is sufficiently high.

5 Conclusion

This paper tried to challenge two puzzles in the welfare benefit program. The first puzzle was 'non-take-up welfare' which means poor people do not take-up welfare even though they are approved to take-up. Second, empirical evidence suggested that there may exist the inverse U-shaped relationship between benefit level and beneficiary ratio. We presented a model of welfare stigma as a hypothesis to explain the above puzzles. Specifically, we investigated the statistical discrimination view model.

When the ratio of elasticity of material utility to benefit level among eligible and non-eligible type decrease in the benefit level, theoretical results are very consistent with empirical evidences. We want to present a model that can explain the situation as future work.

In our model, non-eligible types' decision-makings are extensive. We would like to try to extend our model to consider both intensive and extensive margin.

References

Akerlof, G. A. and R. E. Kranton (2000): "Economics and Identity," *Quarterly Journal of Economics*, 115, 715–753.

Arellano, M. (1987): "Computing Robust Standard Errors for Within-groups Estimators," *Oxford Bulletin of Economics and Statistics*, 49, 431–434.

Baltagi, B. H. (1984): "A Monte Carlo Study for Pooling Time Series of Cross-Section Data in the Simultaneous Equations Model," *International Economic Review*, 25, 603–624.

——— (2008): *Econometric analysis of panel data*, John Wiley & Sons.

Besley, T. and S. Coate (1992): "Understanding welfare stigma: Taxpayer resentment and statistical discrimination," *Journal of Public Economics*, 48, 165–183.

Blank, R. M. and P. Ruggles (1996): "When do women use aid to families with dependent children and food stamps?" *Journal of Human Resources*, 31, 57–89.

Blumkin, T., Y. Margalioth, and E. Sadka (2015): "Welfare Stigma Re-Examined," *Journal of Public Economic Theory*, 17, 874–886.

Currie, J. (2006): "The Take-up of Social Benefits," in *Public Policy and the Income Distribution*, ed. by A. J. Auerbach, D. Card, and J. M. Quigley, New York: Russell Sage Foundation, 80–148.

Duclos, J.-Y. (1995): "Modelling the take-up of state support," *Journal of Public Economics*, 58, 391–415.

Greene, W. H. (2012): *Econometric analysis*, Pearson, 7th ed.

Hausman, J. A. (1978): "Specification Tests in Econometrics," *Econometrica*, 46, 1251–1271.

Honda, Y. (1985): "Testing the Error Components Model with Non-Normal Disturbances," *The Review of Economic Studies*, 52, 681–690.

Hupkau, C. and F. Maniquet (2018): "Identity, non-take-up and welfare conditionality," *Journal of Economic Behavior & Organization*, 147, 13–27.

Immervoll, H. (2009): "Minimum-Income Benefits in OECD Countries: Policy Design, Effectiveness and Challenges," *IZA Discussion Paper No. 4627*.

Kranton, R. (2016): "Identity Economics 2016: Where Do Social Distinctions and Norms Come From?" *American Economic Review: Papers & Proceedings*, 286, 139–146.

Major, B., J. F. Dovidio, and B. G. Link (2018): *The Oxford handbook of stigma, discrimination, and health*, Oxford University Press.

Moffitt, R. (1983): "An economic model of welfare stigma," *American economic review*, 73, 1023–1035.

OECD (2019): "OECD.Stat," *(Accessed on 1 February 2019)*.

Plueger, D. (2009): "Earned Income Tax Credit participation rate for tax year 2005," *IRS Research Bulletin*.

Riphahn, R. T. (2001): "Rational poverty or poor rationality? The take-up of social assistance benefits," *Review of income and wealth*, 47, 379–398.

Tachibanaki, T. and K. Urakawa (2006): *Nihon no Hinkon Kenkyu [A Study of the Japanese Poverty]*, Tokyo: University of Tokyo Press.

Appendix

Proof of Proposition 1. Equilibrium equations are as follows:

$$
\begin{cases}
p = \dfrac{\hat{u}\left(b, z_1\right)}{\phi s\left(\pi, z_1\right)}, \\
q = \dfrac{\hat{u}\left(b, z_2\right)}{\phi s\left(\pi, z_2\right)}, \\
\pi = \dfrac{p}{q}.
\end{cases}
$$

By logarithmic transformation, we obtain the following:

$$
\begin{cases}
\ln p = \ln \hat{u}\left(b, z_1\right) - \ln s\left(\pi, z_1\right) - \ln \overline{\phi}, \\
\ln q = \ln \hat{u}\left(b, z_2\right) - \ln s\left(\pi, z_2\right) - \ln \overline{\phi}, \\
\ln \pi = \ln p - \ln q.
\end{cases}
$$

By totally differentiating and setting $d\theta = dw = d\overline{\phi} = dz_1 = dz_2 = d\gamma = 0$,

$$
\begin{cases}
\dfrac{dp}{p} = \dfrac{\partial \hat{u}\left(b, z_1\right)/\partial b}{\hat{u}\left(b, z_1\right)} db - \dfrac{\partial s\left(\pi, z_1\right)/\partial \pi}{s\left(\pi, z_1\right)} d\pi, \\
\dfrac{dq}{q} = \dfrac{\partial \hat{u}\left(b, z_2\right)/\partial b}{\hat{u}\left(b, z_2\right)} db - \dfrac{\partial s\left(\pi, z_2\right)/\partial \pi}{s\left(\pi, z_2\right)} d\pi, \\
\dfrac{d\pi}{\pi} = \dfrac{dp}{p} - \dfrac{dq}{q}.
\end{cases}
$$

$$\Longleftrightarrow$$

$$
\begin{cases}
\dfrac{dp}{p} = \dfrac{\partial \hat{a}\left(b, z_1\right)}{\partial b} \dfrac{b}{\hat{u}\left(b, z_1\right)} \dfrac{db}{b} - \dfrac{\partial s\left(\pi, z_1\right)}{\partial \pi} \dfrac{\pi}{s\left(\pi, z_1\right)} \dfrac{d\pi}{\pi}, \\
\dfrac{dq}{q} = \dfrac{\partial \hat{a}\left(b, z_2\right)}{\partial b} \dfrac{b}{\hat{u}\left(b, z_2\right)} \dfrac{db}{b} - \dfrac{\partial s\left(\pi, z_2\right)}{\partial \pi} \dfrac{\pi}{s\left(\pi, z_2\right)} \dfrac{d\pi}{\pi}, \\
\dfrac{d\pi}{\pi} = \dfrac{dp}{p} - \dfrac{dq}{q}.
\end{cases}
$$

$$\Longleftrightarrow$$

$$
\begin{cases}
\dfrac{dp}{p} = \eta_b\left(z_1\right) \dfrac{db}{b} + \varepsilon_\pi\left(z_1\right) \dfrac{d\pi}{\pi}, \\
\dfrac{dq}{q} = \eta_b\left(z_2\right) \dfrac{db}{b} + \varepsilon_\pi\left(z_2\right) \dfrac{d\pi}{\pi}, \\
\dfrac{d\pi}{\pi} = \dfrac{dp}{p} - \dfrac{dq}{q}.
\end{cases}
$$

A matrix representation is given below:

$$\begin{bmatrix} 1 & 0 & -\varepsilon_\pi(z_1) \\ 0 & 1 & -\varepsilon_\pi(z_2) \\ 1 & -1 & -1 \end{bmatrix} \begin{bmatrix} \frac{dp/p}{db/b} \\ \frac{dq/q}{db/b} \\ \frac{d\pi/\pi}{db/b} \end{bmatrix} = \begin{bmatrix} \eta_b(z_1) \\ \eta_b(z_2) \\ 0 \end{bmatrix}.$$

By Cramer's rule, solutions are given as follows:

$$\frac{dp/p}{db/b} = \frac{-\eta_b(z_1)[1+\varepsilon_\pi(z_2)] + \eta_b(z_2)\varepsilon_\pi(z_1)}{\varepsilon_\pi(z_1) - [1+\varepsilon_\pi(z_2)]},$$
$$\frac{dq/q}{db/b} = \frac{\eta_b(z_2)[1+\varepsilon_\pi(z_1)] - \eta_b(z_1)\varepsilon_\pi(z_2)}{\varepsilon_\pi(z_1) - [1+\varepsilon_\pi(z_2)]},$$
$$\frac{d\pi/\pi}{db/b} = \frac{-\eta_b(z_1) + \eta_b(z_2)}{\varepsilon_\pi(z_1) - [1+\varepsilon_\pi(z_2)]}.$$

Since the stability condition is $\varepsilon_{\pi^*}(z_1) - \varepsilon_{\pi^*}(z_2) < 1$, the denominator, $\varepsilon_\pi(z_1) - [1+\varepsilon_\pi(z_2)]$, is negative.

Therefore, the result of comparative statics regarding a change in benefit level is given as follows:

$$\operatorname{sgn}\left[\frac{dp^*}{db}\right] = \operatorname{sgn}\left[\frac{\eta_b(z_1)}{\eta_b(z_2)} - \frac{\varepsilon_{\pi^*}(z_1)}{1+\varepsilon_{\pi^*}(z_2)}\right],$$
$$\operatorname{sgn}\left[\frac{dq^*}{db}\right] = \operatorname{sgn}\left[\frac{\eta_b(z_1)}{\eta_b(z_2)} - \frac{1+\varepsilon_{\pi^*}(z_1)}{\varepsilon_{\pi^*}(z_2)}\right],$$
$$\operatorname{sgn}\left[\frac{d\pi^*}{db}\right] = \operatorname{sgn}\left[\frac{\eta_b(z_1)}{\eta_b(z_2)} - 1\right].$$

■

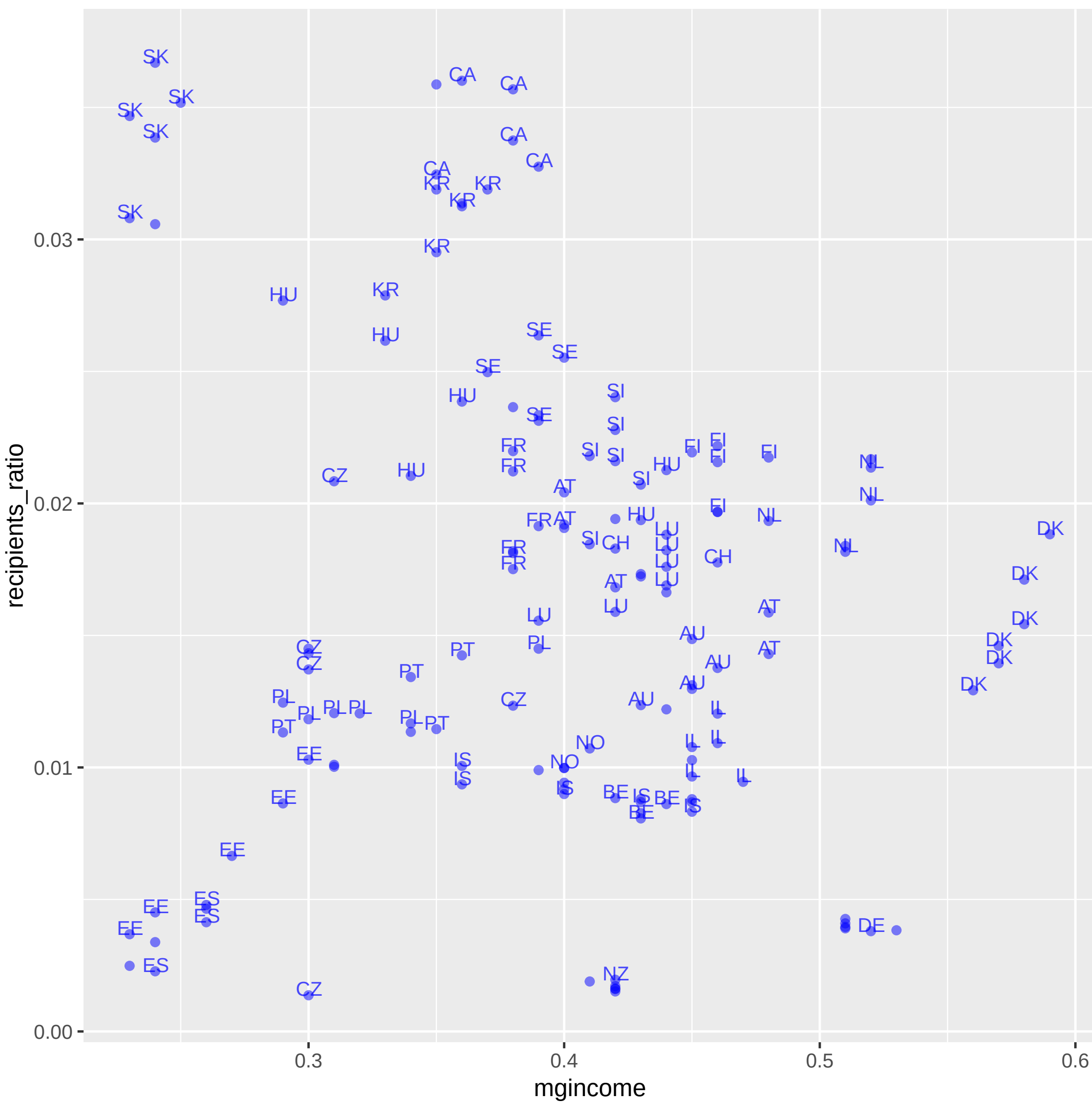

Figure 1: Simple relationship between beneficiary ratio and benefit level

Notes: Strings accompanied by points indicate ISO 3166-1 alpha-2 code of countries.

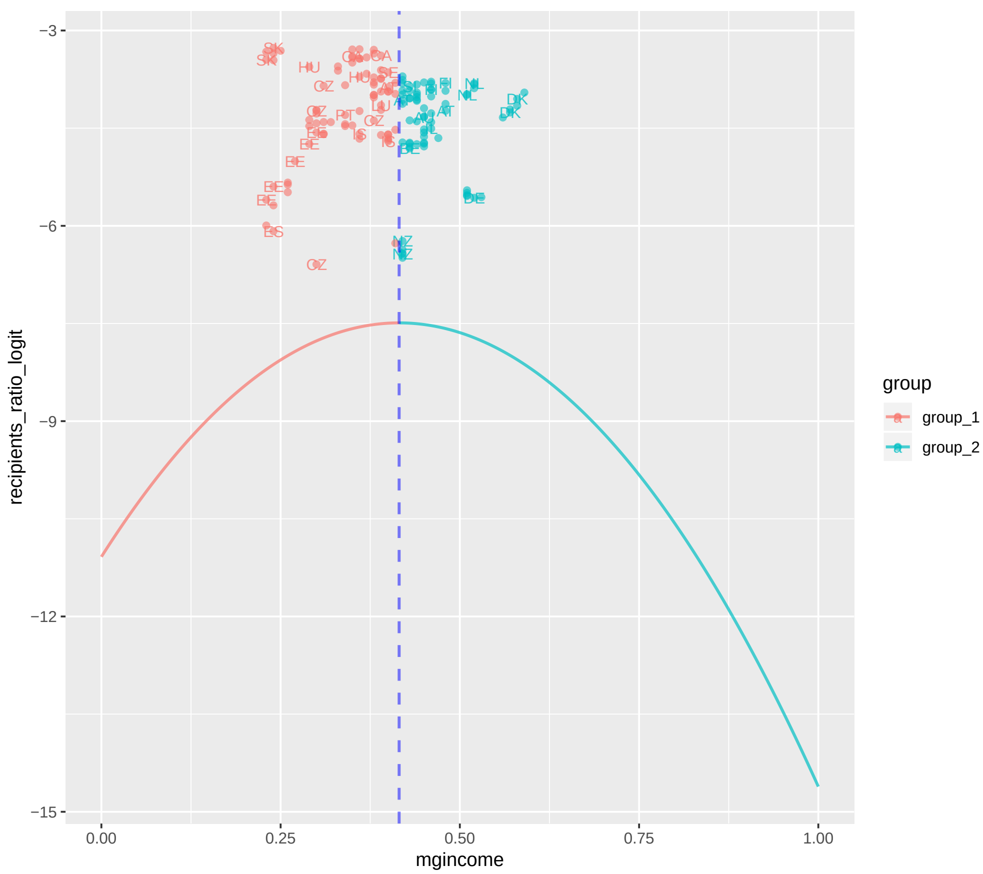

Figure 2: Fitted curve of one-way RE and scatter plot between beneficiary ratio and benefit level

Notes: Strings accompanied by points indicate ISO 3166-1 alpha-2 code of countries. The dashed blue line indicates a location of `mgincome` which corresponds to an estimated maximum of `recipients_ratio_logit` obtained by the fitted curve of one-way RE. `group_1` and `group_2` are defined by the location based on the dashed blue line. If an observed value of `mgincome` is less than the dashed blue line, it is categorized as `group_1` otherwise it is categorized as `group_2`.

Table 1: Descriptive stastistics of OECD panel data: whole data

	Mean	Median	Standard Deviation	Min	Max
recipients_ratio	0.016	0.015	0.009	0.001	0.037
logit_recipients_ratio	-4.327	-4.175	0.753	-6.591	-3.268
mgincome	0.397	0.400	0.082	0.230	0.590
mgincome_2	0.164	0.160	0.065	0.053	0.348
gdp_capita	37704.939	37699.559	14081.170	16788.433	91814.013
log_gdp_capita	10.479	10.537	0.335	9.728	11.428
youth_dependency	0.254	0.241	0.053	0.199	0.459
old_dependency	0.229	0.238	0.040	0.138	0.314
divorce_rate	0.002	0.002	0.000	0.001	0.003
unemployment	0.074	0.072	0.037	0.023	0.249

Notes: $T = 6, n = 25, N = 150$.

Table 2: Descriptive stastistics of OECD panel data: by country

country	recipients_ratio	logit_recipients_ratio	mgincome	mgincome_2	gdp_capita	log_gdp_capita	youth_dependency	old_dependency	divorce_rate	unemployment
					Mean (Standard Deviation)					
Australia	0.013 (0.001)	-4.315 (0.074)	0.447 (0.010)	0.200 (0.009)	40852.112 (2156.481)	10.617 (0.053)	0.284 (0.002)	0.201 (0.007)	0.002 (0.000)	0.050 (0.005)
Austria	0.018 (0.002)	-4.029 (0.138)	0.430 (0.039)	0.186 (0.035)	42435.028 (2584.710)	10.654 (0.060)	0.222 (0.006)	0.259 (0.005)	0.002 (0.000)	0.048 (0.004)
Belgium	0.009 (0.000)	-4.744 (0.035)	0.437 (0.012)	0.191 (0.011)	39517.464 (2220.265)	10.583 (0.056)	0.257 (0.001)	0.261 (0.003)	0.003 (0.000)	0.076 (0.005)
Canada	0.034 (0.002)	-3.335 (0.049)	0.368 (0.017)	0.136 (0.013)	40428.026 (1271.527)	10.607 (0.031)	0.239 (0.004)	0.203 (0.008)	0.002 (0.000)	0.073 (0.010)
Czech Republic	0.013 (0.006)	-4.592 (0.996)	0.315 (0.032)	0.100 (0.022)	27776.316 (1070.706)	10.231 (0.039)	0.205 (0.006)	0.218 (0.014)	0.003 (0.000)	0.063 (0.011)
Denmark	0.015 (0.002)	-4.161 (0.140)	0.575 (0.010)	0.331 (0.012)	42137.252 (2322.734)	10.647 (0.055)	0.276 (0.004)	0.251 (0.014)	0.003 (0.000)	0.061 (0.020)
Estonia	0.007 (0.003)	-4.984 (0.433)	0.273 (0.033)	0.076 (0.018)	22846.807 (2032.578)	10.033 (0.088)	0.226 (0.006)	0.259 (0.005)	0.002 (0.000)	0.107 (0.048)
Finland	0.021 (0.001)	-3.837 (0.056)	0.462 (0.010)	0.213 (0.009)	39257.144 (1350.605)	10.577 (0.034)	0.252 (0.002)	0.261 (0.014)	0.002 (0.000)	0.076 (0.008)
France	0.019 (0.002)	-3.929 (0.094)	0.382 (0.004)	0.146 (0.003)	35822.205 (1475.914)	10.486 (0.041)	0.284 (0.003)	0.261 (0.007)	0.002 (0.000)	0.085 (0.009)
Germany	0.004 (0.000)	-5.525 (0.044)	0.515 (0.008)	0.265 (0.009)	39925.940 (2678.832)	10.593 (0.066)	0.205 (0.002)	0.310 (0.005)	0.002 (0.000)	0.071 (0.012)
Hungary	0.023 (0.003)	-3.747 (0.142)	0.365 (0.059)	0.136 (0.044)	21298.496 (1512.043)	9.964 (0.072)	0.215 (0.003)	0.241 (0.006)	0.002 (0.000)	0.098 (0.017)
Iceland	0.009 (0.001)	-4.710 (0.076)	0.405 (0.038)	0.165 (0.031)	41380.192 (1215.725)	10.630 (0.029)	0.311 (0.003)	0.179 (0.008)	0.002 (0.000)	0.056 (0.024)
Israel	0.011 (0.001)	-4.547 (0.090)	0.457 (0.008)	0.209 (0.007)	28872.669 (1899.220)	10.269 (0.065)	0.454 (0.005)	0.160 (0.004)	0.002 (0.000)	0.083 (0.011)
Korea	0.031 (0.002)	-3.456 (0.055)	0.353 (0.014)	0.125 (0.010)	29748.950 (1733.041)	10.299 (0.058)	0.227 (0.016)	0.149 (0.009)	0.002 (0.000)	0.035 (0.002)
Luxembourg	0.017 (0.001)	-4.050 (0.077)	0.428 (0.020)	0.184 (0.017)	86918.694 (3971.080)	11.372 (0.045)	0.259 (0.008)	0.204 (0.002)	0.002 (0.000)	0.048 (0.005)
Netherlands	0.020 (0.001)	-3.902 (0.076)	0.510 (0.015)	0.260 (0.015)	45641.618 (1321.341)	10.728 (0.029)	0.263 (0.003)	0.230 (0.012)	0.002 (0.000)	0.041 (0.008)
New Zealand	0.002 (0.000)	-6.368 (0.100)	0.418 (0.004)	0.175 (0.003)	30845.995 (1511.117)	10.336 (0.049)	0.316 (0.001)	0.195 (0.008)	0.002 (0.000)	0.058 (0.014)
Norway	0.010 (0.001)	-4.610 (0.054)	0.400 (0.006)	0.160 (0.005)	59776.358 (3956.631)	10.997 (0.066)	0.285 (0.004)	0.226 (0.006)	0.002 (0.000)	0.031 (0.004)
Poland	0.012 (0.001)	-4.378 (0.081)	0.325 (0.036)	0.107 (0.025)	20207.247 (2579.601)	9.907 (0.129)	0.215 (0.003)	0.191 (0.003)	0.002 (0.000)	0.092 (0.012)
Portugal	0.012 (0.002)	-4.419 (0.128)	0.332 (0.026)	0.111 (0.017)	26558.193 (521.970)	10.187 (0.020)	0.230 (0.004)	0.277 (0.010)	0.002 (0.000)	0.113 (0.032)
Slovak Republic	0.034 (0.002)	-3.360 (0.076)	0.238 (0.008)	0.057 (0.004)	24142.536 (2038.476)	10.089 (0.086)	0.217 (0.004)	0.173 (0.005)	0.002 (0.000)	0.125 (0.019)
Slovenia	0.022 (0.002)	-3.818 (0.092)	0.418 (0.008)	0.175 (0.006)	28351.080 (880.111)	10.252 (0.031)	0.203 (0.004)	0.238 (0.006)	0.001 (0.000)	0.067 (0.018)
Spain	0.004 (0.001)	-5.658 (0.320)	0.248 (0.013)	0.062 (0.007)	32403.199 (568.189)	10.386 (0.017)	0.218 (0.005)	0.247 (0.008)	0.002 (0.000)	0.173 (0.063)
Sweden	0.024 (0.001)	-3.686 (0.055)	0.387 (0.010)	0.150 (0.008)	42080.770 (1911.048)	10.646 (0.045)	0.257 (0.003)	0.279 (0.012)	0.002 (0.000)	0.076 (0.011)
Switzerland	0.018 (0.001)	-4.013 (0.055)	0.433 (0.015)	0.188 (0.013)	53399.177 (3100.762)	10.884 (0.058)	0.220 (0.007)	0.252 (0.011)	0.002 (0.000)	0.042 (0.006)

Notes: $T = 6, n = 25, N = 150$Numbers in parentheses stand for the standard deviation.

Table 3: Descriptive stastistics of OECD panel data: by year

year	recipients_ratio	logit_recipients_ratio	mgincome	mgincome_2	gdp_capita	log_gdp_capita	youth_dependency	old_dependency	divorce_rate	unemployment
					Mean					
2007	0.016	-4.339	0.396	0.164	35645.300	10.419	0.257	0.221	0.002	0.060
2008	0.015	-4.383	0.391	0.160	37381.672	10.468	0.255	0.223	0.002	0.057
2009	0.016	-4.311	0.397	0.164	36290.570	10.443	0.253	0.226	0.002	0.078
2010	0.017	-4.275	0.400	0.166	37469.973	10.476	0.252	0.230	0.002	0.084
2011	0.016	-4.370	0.399	0.166	39355.518	10.524	0.252	0.234	0.002	0.081
2012	0.016	-4.284	0.398	0.166	40086.599	10.543	0.252	0.240	0.002	0.083
					Median					
2007	0.016	-4.147	0.400	0.160	36871.534	10.515	0.249	0.229	0.002	0.054
2008	0.014	-4.231	0.400	0.160	38133.413	10.549	0.242	0.234	0.002	0.056
2009	0.015	-4.212	0.400	0.160	37695.802	10.537	0.240	0.237	0.002	0.078
2010	0.015	-4.156	0.400	0.160	38737.069	10.565	0.237	0.239	0.002	0.077
2011	0.015	-4.194	0.420	0.176	40683.337	10.614	0.236	0.240	0.002	0.072
2012	0.016	-4.127	0.420	0.176	40619.937	10.612	0.236	0.248	0.002	0.074
					Standard Deviation					
2007	0.009	0.775	0.082	0.064	13813.492	0.353	0.054	0.039	0.000	0.024
2008	0.008	0.758	0.084	0.064	14436.887	0.348	0.054	0.040	0.000	0.022
2009	0.009	0.731	0.082	0.065	13382.286	0.336	0.053	0.041	0.000	0.033
2010	0.009	0.728	0.080	0.065	13804.461	0.330	0.053	0.041	0.000	0.040
2011	0.010	0.865	0.084	0.068	14869.727	0.330	0.053	0.041	0.000	0.040
2012	0.009	0.726	0.089	0.071	15019.149	0.329	0.054	0.042	0.000	0.046

Notes: $T = 6, n = 25, N = 150$.

Table 4: Results of empirical analysis using OECD panel data

	Dependent variable:				
	logit_recipients_ratio				
	Pooling	one-way FE	two-way FE	one-way RE	two-way RE
mgincome	15.277**	18.192***	19.529***	17.301***	18.596***
	(6.340)	(6.297)	(6.327)	(5.471)	(5.537)
mgincome_2	−16.284**	−22.255***	−23.934***	−20.831***	−22.347***
	(7.509)	(8.200)	(8.239)	(7.116)	(7.193)
log_gdp_capita	0.206	0.553	0.465	0.518	0.490
	(0.214)	(0.491)	(0.739)	(0.330)	(0.397)
unemployment	2.763	4.845***	5.215***	4.730***	4.951***
	(2.378)	(1.270)	(1.476)	(1.166)	(1.328)
youth_dependency	−6.583***	−5.926	−7.378	−5.629**	−6.458**
	(1.458)	(4.298)	(4.665)	(2.570)	(2.639)
old_dependency	−7.790***	−6.112*	−9.749*	−6.518**	−8.839***
	(2.059)	(3.463)	(5.165)	(2.688)	(3.261)
divorce_rate	76.239	240.823	275.167	206.935	250.086
	(171.190)	(166.441)	(170.419)	(146.998)	(155.897)
Constant	−6.789***			−11.083***	−10.424**
	(2.436)			(3.619)	(4.343)
Observations	150	150	150	150	150
R^2	0.163	0.209	0.220	0.196	0.229
Adjusted R^2	0.121	0.001	−0.029	0.156	0.161
F-test (vs. pooling)		55.834 ***	46.667***		
F-test (vs. one-way FE)			1.1348		
LM-test (vs. pooling)				17.320***	11.152***
Hausman-test (vs. random effect)		0.96729	0.65122		

Notes: Numbers in parentheses stand for standard error calculated by HAC (Arellano, 1987) estimator. Above *,** ,*** indicate statistical significance at 10%, 5%, 1%, respectively.