Essays on Economic Models of Welfare Stigma

栗田, 健一

https://doi.org/10.15017/2534373

出版情報：Kyushu University, 2019, 博士（経済学）, 課程博士
バージョン：
権利関係：

Kyushu University

Doctoral Thesis

Essays on Economic Models of Welfare Stigma

Author:

Kenichi Kurita

Supervisor:

Prof. Nobuaki Hori

Co-supervisor:

Prof. Kunio Urakawa

Prof. Tetsushi Murao

A thesis submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Economics

in the

Department of Economic Engineering

Graduate School of Economics

July, 2019

Abstract

Kenichi KURITA

Essays on Economic Models of Welfare Stigma

This thesis presents a collection of attempting to examine the relationship of the interaction between welfare stigma and household decision-making with respect to participating in welfare policy by analysing theoretical economic models from various views.

In the first chapter, I present the background, the motivation and the survey of the related literature. I explain standard labour supply model and previous researches on welfare stigma. Then, I suggest their limitations to explain actual states in the real world.

In the second chapter, I analyse the extended version of statistical discrimination stigma model. The model was developed by Besley and Coate (1992). There exist two types in the model: the needy type and the non-needy type. The needy type people cannot work even if they hope to do. On the other hand, the non-needy type people can work even if they want to do. Stigma cost is determined by the ratio of the non-needy type to the needy type in recipients. In particular, they assumed that stigma cost is a decreasing function with the ratio. Their results indicated the occurrence of welfare fraud. However, the needy type was assumed to take-up welfare regardless of level of stigma cost. In general, people like needy type are thought to be influenced by stigma in their

decision-makings. Moreover, not taking-up welfare by needy type corresponds to 'ro-kyu' which means not taking-up welfare by eligible poor people. To solve these problems and limitations in previous research, I extend the statistical stigma model to endogenise the decision-making process for the needy type. As a result, multiple equilibrium can occur: one equilibrium is 'high stigma, serious ro-kyu and restrained welfare fraud' and the other is 'low stigma, restrained ro-kyu, serious welfare fraud'. Multiple equilibria are likely to occur when the needy type's elasticity of stigma sensitiveness to the ratio of the needy type to the non-needy type in the pool of recipients in welfare benefit. The comparative static analysis indicates that an increase in the benefit level is likely to reduce take-up of welfare benefits in needy types. I demonstrate this result through a simple equation, linking two types of elasticity for needy and non-needy types: the needy type's elasticity of stigma sensitiveness to the ratio between types in recipients and the non-needy type's elasticity of material utility to benefit level are sufficiently large.

The third chapter studies the stigma model of relative income. The model is similar to the taxpayer resentment model (Besley and Coate, 1992) in the point that non-recipients (workers, taxpayers or capitalists). The model differs from Besley and Coate (1992)'s taxpayer resentment model in that the stigmatisee can become the stigmatiser. Moreover, we use relative income mechanism to formulate stigma which is increasing function with average of the ratio of benefit level to each working wage. As a result, multiple equilibria can occur, high stigma and low stigma, contrary to the proportion of Besley and Coate (1992)'s model. It is because there exists feedback effect. In the comparative static analysis, the result indicates that the possibility that the number of recipients declines in case of a negative macroeconomic shock since there exists the negative indirect effect on recipients, which decreases in case of increase in stigma from increased resentment due to a shift of working income distribution.

The fourth chapter researches the aspiration enhancing role of the stigma created by low income and poverty. I discuss the possibility that this stigma gives children an incentive to seek an escape from poverty, and I investigate the mechanism of transmitting this stigma culture over generations. In some cases, increased income inequality immobilises family cultural capital, and social mobility is constrained.

The fifth chapter investigates the relationship between benefit level and the beneficiary ratio in minimum income guarantee program from the view of welfare stigma using theoretical analysis and empirical evidence. First, the theoretical study presents a simple stigma model to consider household decision-making regarding whether to work or take up welfare benefits. As a theoretical result, an equilibrium recipient ratio forms the inverse U-shaped curve with respect to benefit level when stigma cost is an increasing convex function with benefit level. On the other hand, an equilibrium recipient ratio appears to be the U-shaped function with benefit level when stigma cost is an increasing concave function with benefit level. Second, an empirical analysis is conducted using OECD panel data to examine which case of stigma is in keeping with the estimation result regarding the relationship between benefit level and beneficiary ratio. The empirical results are consistent with the case that stigma cost is convex function with respect to benefit level.

The final chapter concludes this thesis.

Acknowledgements

First, I am grateful to my main adviser, Professor Nobuaki Hori. Thank You, Professor Hori.

Next, I would like to thank my dissertation committee members, Professor Kunio Urakawa and Professor Tetsushi Murao for their insightful suggestions and helpful supports.

I could learned a lot of things through collaborative research. In particular, Professor Jun-ichi Itaya helped me evolve through co-research on evolutionary games. That experience has been food for my research and life. Thank you, Professor Itaya. I would like to express appreciation for Professor Yuya Katafuchi, you are my best friend and the great teacher of empirical analysis for me. Thank you, Yuya. Professor Yuji Tamura has given me important comments my research and told me how to think as the economist through joint study. Thank you, Professor Tamura.

I have been helped by comments from many people at workshops, seminars, conference or other conversations, even though I cannot express my gratitude to all of them. I would like to thank Professor Keisuke Hattori, Professor Takeshi Yoshikawa, Professor Yusuke Zennyo, Professor Kosuke Hirose, Professor Jumpei Hamamura, Professor Kimiyuki Morita, Professor Shuichi Tsugawa, Professor Yoshifumi Kon, Professor Tohru Naito, Professor Isao Miura, Professor Masayuki Kanazaki, Professor Takeshi Miyazaki, Professor Norikazu Tawara, Professor Kimiko Terai, Professor Takeo Hori, Professor Hirokazu Takizawa, Professor Susumu Kato, Professor Ryoichi Imai, Professor Hiroki Arato, Professor Kengo Nutahara, Professor Shiba Suzuki.

Finally, I would like to thank my family for their encouragement.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis presents an economic analysis of the welfare stigma to understand the impact of welfare programmes to alleviate poverty. Many economists analyse welfare programmes using the time allocation model from Becker (1965) (Moffitt, 1983; Atkinson, 1995). However, the model cannot explain certain phenomena that occurs in the real world. For example, two important problems exist in welfare programmes: welfare fraud and non-take-up welfare. Welfare fraud means taking-up welfare by non-deserving people. On the other hand, non-take-up welfare corresponds to not taking-up welfare by deserving poor people. Welfare fraud can occur in the standard model but non-take-up welfare cannot. In other words, certain factors are not considered in the model. One such factor is welfare stigma. Stigma is sociological concept describing a negative label applied to behaviour by society or a social group. In particular, stigma is an important concept in social psychology and is researched in the context of stereotype or identity (Major, Dovidio, and Link, 2017). Welfare stigma is related to welfare policy. The existence of stigma is suggested in studies on economics (Moffitt, 1983; Kayser and Frick, 2000) and sociology (Horan and Austin, 1974; Rainwater, 1982).

This paper presents a collection of attempting to examine the relationship of the interaction between stigma and household decision-making by analysing theoretical economic models from various views.

The next chapter presents a model of statistical discrimination view of stigma to consider decision-making between workable and non-workable poor people. This model is the extended version of the statistical discrimination view stigma model of Besley and Coate (1992)'s. As a result, in Besley and Coate (1992), welfare fraud occurs but non-take-up welfare cannot occur. This result is because they assumed that the utility function from consumption is minus infinity at the zero consumption level; that is, the assumption is equal to a non-workable type always claiming to take-up welfare.

The third chapter presents the relative income stigma model. In this model, I use a new mechanism to constitute stigma. The mechanism is based on the notion that workers feel anger over welfare recipients' take-up of benefits without working. Such anger is higher when the ratio of benefit level to worker's wage is higher. A comparative static analysis shows an interesting result: recipients' equilibrium level declines when negative macroeconomic shocks occur.

The fifth chapter presents the cultural transmission model. I apply the Bisin and Verdier (2000) and Bisin and Verdier (2001)' cultural transmission model to consider an endogenously inter-generational formation of a value system to feel the stigma.

The sixth chapter presents the research that investigates the relationship between benefit level and the beneficiary ratio in minimum income guarantee programmes from the viewpoint of welfare stigma using theoretical analysis and empirical evidence. First, the theoretical section analyses a simple stigma model to consider household decision- making regarding whether to work or take-up welfare benefits. Second, an empirical analysis is conducted using OECD panel data to examine which case of stigma that is in keeping with the estimation result regarding the relationship between benefit level and beneficiary ratio. The empirical results are consistent with that the cost of stigma is a convex function in the theoretical model with respect to benefit level. The final chapter concludes this thesis.

1.1 Survey of theoretical researches on welfare stigma

> *"Welfare stigma is widely regarded to be of central importance in understanding the impact of welfare programmes in their goal of alleviating poverty."* (P. 165, Besley and Coate, 1992)

Welfare policy is considered an important policy in the fields of public economics, labour economics and other applied microeconomics. For example, in most textbooks on microeconomics or applied fields, minimum income guarantee problem and EITC are analysed using standard models and are suggested for some problems. Moreover, in research on economics, poverty, inequality and welfare are interesting to many people. In particular, the basic income programme is argued in many countries in connection with the next industrial revolution of artificial intelligence.

Most economists use the standard labour supply model based on the maximisation problem regarding time allocation to analyse welfare programmes (Becker, 1965). That model points out the important result that incentive problems exist in minimum guarantee programmes as observed in the next section. To solve this problem, many economists argued over alternative policies (Friedman, 1962; Atkinson, 1995; Fitzpatrick, 1999; Moffitt, 2003). However, these researchers focused on welfare fraud but not non-take-up welfare. Moreover, the standard model can explain distortions in labour incentives or welfare fraud but cannot explain the occurrence of non-take-up welfare.

This chapter presents a review of research on welfare policy and some important practical cases in welfare programmes. The structure of this survey is organised as follows. Next, I see Becker's standard model based on the maximisation problem in two goods; a leisure and a consumption good. In this section, I confirm important implications and limitations in analysing welfare benefit programmes by using a standard model. The second section presents a review of Besley and Coate (1992) which presented the model in which the level of stigma cost is endogenous. The final section presents a review of empirical research on welfare stigma.

1.1.1 Standard labour supply model

As the benchmark, the standard labour supply model of welfare benefit is important. This model presented an application of the optimal allocation of time (Becker, 1965). There are two goods; leisure, L, consumption good, C. Here, the price of C is normalised to one. The maximization problem is formulated as follows:

$$
\begin{aligned}
&\max_{L,C} U(L,C) \\
\text{s.t.}\quad & L + H \leq T, \\
& C \leq wH + B + I, \\
& B = \max[0, G - wH],
\end{aligned}
$$

where H is working hours, T is available times, B is benefit level, I is non-wage income, G is the minimum income guarantee level, and w is the hourly wage. The result in this model is summarised in the following remark. The type of welfare benefit program in this setting called minimum income guarantee program.

Proposition 1 *In standard labour supply model (Becker, 1965), the optimal labour supply in minimum income guarantee program is given as follows*

$$H^* \begin{cases} = T & \text{if} \quad \mathrm{MRS}_{LC}(0, wT + I) < w, \\ \in (0, T) & \text{if} \quad \mathrm{MRS}_{LC}(0, G) < w < \mathrm{MRS}_{LC}(0, wT + I), \\ 0 & \text{otherwise.} \end{cases}$$

In particular, if $\mathrm{MRS}_{LC}(0, G) < w < \mathrm{MRS}_{LC}(0, wT + I)$*,* H^* *is the level which satisfies* $\mathrm{MRS}_{LC}(T - H^*, wH^* + I)$*.*

In the case, $H^* = T$, the household puts all times into labour but enjoy no time for leisure. It is because that the household's marginal rate of substitution of leisure to consumption is sufficiently low. Such an individual is called "workaholic". In the strictly interior case, $H^* \in (0, T)$, the household supplies positive labour and enjoy positive leisure. In the weak interior case, $H^* = 0$, the household does not any supply labour and enjoy all times for leisure.

Many economists focus on and problematise the third case since there can occurs the distortion of incentive in labour supply. To settle this disincentive problem, many alternative programs are presented. Friedman (1962) advocated "negative income taxation". He states there are merits in negative income taxation as follows: the first labour incentive is promoted, the second it can be more directly mechanism rather than other tagging mechanism, the third the cash transfer rather than the Food Stamp Program, the fourth unified welfare policy can eliminate duplication of existing programs, and the administrative cost is lower than the one in existing policies. However, Friedman did not consider clearly the government budget constraint to finance the cost for the program.

Atkinson (1995) presented "the Basic Income/Flat Tax proposal" which is the mechanism combining basic income concept with flat taxation. Atkinson revised Friedman's concept to consider government budget constraint by using flat taxation. In this proposal, the distortion in labour supply is reduced more than the one in the minimum income guarantee problem. However, it does not mean that disincentive vanishes since low income workers reduce each labour supply. In empirical researches, Robins (1985) suggested that aggregate labour supply decreases by implementing negative income taxation in New Jersey (1968 - 1972), Rural Iowa / North Carolina (1969 - 1973), Gary (1971 - 1974), and Settle-Denver (1971-82). In fact, labour time reduced in 5 $\sim$ 25% and employment rate reduces in 1 $\sim$ 10%. It is generated by the income effect from generous transfer and the substitution effect from high marginal tax rate. Moreover, negative income taxation needs large costs in general. Ghatak and Maniquet (2019) present some theoretical aspects with respect to a universal basic income and review the possible justifications of introducing that.

The standard model cannot explain the occurrence of non-take-up welfare. That is, The model does not consider important factors. Welfare stigma is one of such factors. Next section introduces Besley and Coate (1992)'s paper analysing the model of welfare stigma.

1.1.2 Besley and Coate (1992)

Besley and Coate (1992) present the model where the cost of stigma is endogenised. Moreover, they analyse two mechanisms of formation in stigma; the statistical discrimination view stigma and tax-payer resentment view stigma.

Basic setting

There is an economy with two classes: the rich and the poor. Total population is N, the number of the poor is n. There are two types in the poor: the needy and the non-needy. The former type cannot work. The latter can work. A fraction γ of the poor is the needy type.

The rich cares about the level of consumption of the needy poor. Let denote the consumption level of the needy poor as c_n. The rich's utility is assumed as follows:

$$u(y - T) - \mu n \gamma P(c_n),$$

where is a parameter corresponding to a weight representing the individual's degree of concern about the needy poor, and $P(\cdot)$ is a measure of poverty or distress, $P'(\cdot) < 0$, $P^{''}(\cdot) \geq 0$. We assume that μ is distributed among the population with probability distribution function $G(\mu)$. T is a tax and y is an income of the rich.

The welfare program is financed by a flat taxation on the rich.

$$T = mb/(N - n)$$

where m is the total number of poor on welfare, b is a level of welfare benefit.

The non-needy poor who choose to work obtain utility:

$$v(w) - \theta,$$

where w is working wage rate and θ is the disutility of labour. We assume that θ is varies over poor individuals with uniformly distribution function on the interval $[0, 1]$, the mean of $\bar{\theta} := E[\theta] = 1/2$. The function $v(\cdot)$ represents the material utility from income, $v^{'}(\cdot) > 0$, $v^{''}(\cdot) \leq 0$, $v(0) = -\infty$. The last assumption means that some positive level of income is necessary to survive.

Poor individuals who do not work and claim to take-up welfare obtain utility:

$$v(b) - s,$$

where s is an index of stigma. This formulation assumes that all poor individuals feel the same level of the psychic cost of being on welfare.

Since needy poor cannot earn any working income and we assume $v(0) = -\infty$, they always choose to go on welfare. In the other hand, non0needy poor people have two choices: work or go on welfare. We suppose that the government can observe each working income but cannot identify which the claimant is the needy type or the non-needy type. Then, the non-needy can take-up welfare if the individual claim it in this model. To consider the decision-makings in the non-needy poor, I define a critical level of θ, $\hat{\theta}$:

$$v(w) - \hat{\theta} = v(b) - s$$
$$\hat{\theta} = v(w) - v(b) + s$$

Then, all those individuals for whom $\theta \geq \hat{\theta}$ choose to go on welfare. Total number of poor on welfare is given as

$$m = n\{\gamma + (1-\gamma)(1-\hat{\theta})\}$$
$$= n\{\gamma + (1-\gamma)(1 - v(w) + v(b) - s)\}$$

Here are two types of recipients, the one is the "deserving recipients" and the other is "undeserving recipients". Clearly,

$$\frac{\partial m}{\partial s} = -n(1-\gamma) < 0,$$

$$\frac{\partial m}{\partial b} = n(1-\gamma)v^{'}(b) > 0,$$

$$\frac{\partial m}{\partial y} = 0,$$

$$\frac{\partial m}{\partial w} = -n(1-\gamma)v^{'}(w) < 0,$$

$$\frac{\partial m}{\partial \gamma} = n\{v(w) - v(b) + s\} > 0,$$

$$\frac{\partial m}{\partial n} = \{\gamma + (1-\gamma)(1 - v(w) + v(b) - s)\} > 0.$$

Although above effects are very intuitive, they are just direct effect.

In the next two sub-subsections, I formulate the determining of stigma.

The statistical discrimination view stigma

Besley and Coate (1992) suppose the statistical discrimination view stigma inspired the work of American Sociologist, Goffman (1963). Specifically, this stigma is modelled as follows:

$$s = g(\bar{\theta}_w(s) - \bar{\theta}),$$

where

$$g'(\cdot) > 0, g(0) = 0,$$

$$\bar{\theta}_w = \pi\bar{\theta} + (1-\pi)\bar{\theta}_u,$$

$$\bar{\theta}_u = \int_{\hat{\theta}}^{1} \frac{\theta}{1-\hat{\theta}} d\theta = \frac{1+\hat{\theta}}{2},$$

$$\pi = \frac{n\gamma}{m} = \frac{\gamma}{\{\gamma + (1-\gamma)(1-\hat{\theta})\}}.$$

The function $g(\cdot)$ maps the differences in the disutility of labour into the index of stigma, and $g(\cdot)$ is an increasing function with the difference.

The equilibrium level of welfare stigma is satisfies the equation:

$$s^* = g(\bar{\theta}_w(s^*) - \bar{\theta}).$$

Therefore, the equilibrium corresponds to the fixed point in self-mapping.

Besley and Coate (1992) conduct comparative static analysis regarding effects on welfare stigma of change in parameters. Results are summarised in the following proposition.

Proposition 2 *Besley and Coate (1992): Comparative statics in the statistical discrimination view stigma:*

- *An increase in benefits increases welfare stigma if* $\hat{\theta} > \frac{1-\sqrt{\gamma}}{1-\gamma}$.
- *Changes in the income of the rich and changes in the number of poor do not affect welfare stigma.*
- *An increase in the wage of the non-needy poor decreases welfare stigma* $\hat{\theta} > \frac{1-\sqrt{\gamma}}{1-\gamma}$.
- *An increase in the fraction of the poor who are needy decreases welfare stigma.*

The sign of an effect on welfare stigma of an increase in benefit level is obscure. It is because there exist two effects, the direct effect and the indirect effect. The direct effect is positive and generated by that the fraction of undeserving claimants increases with benefit level. The indirect effect is negative and driven by that higher level of welfare gives incentive regarding "take-up welfare" to the non-needy poor who have lower disutilities of labour on average. Because of such oppositional effects, the sign of a change in stigma with respect to an increase in benefit level is ambiguous. By simplifying equation of $\frac{\partial \bar{\theta}_w}{\partial b}$, I can attribute the determinant of the sign to the relationship between γ and $\hat{\theta}$. That is to say, the sign is positive if $\hat{\theta}$ is greater than $\frac{1-\sqrt{\gamma}}{1-\gamma}$. Fixed γ, lower benefit level, the sign of $\frac{\partial \bar{\theta}_w}{\partial b}$ is more likely to get positive.

The sign of an effect on stigma of an increase in wage of the working poor is unclear, too. It can be considered as a reduction in benefit level. Therefore, the sign is negative if $\hat{\theta}$ is greater than $\frac{1-\sqrt{\gamma}}{1-\gamma}$. Fixed γ, lower wage, the sign of $\frac{\partial \bar{\theta}_w}{\partial b}$ is more likely to get positive.

The sign of an effect on stigma of an rise in proportion of the needy type in the poor is negative. It is caused by that the fraction of the deserved claimants increases with the proportion, γ.

The taxpayer resentment view stigma

The optimal benefit level for the rich who has μ is given as

$$b^* = \arg\max \left[u\left(y - \frac{mb}{N-n} \right) - \mu n \gamma P(b) \right]$$

Besley and Coate (1992) suppose that the taxpayers whom μ is less than the following $\hat{\mu}$

$$b^*(\hat{\mu}, s) = b$$

regard the level of welfare benefit as "too generous", and stigma is generated by their resentments to welfare claimants. Let $r(\mu, s, b)$ denote as the index of the resentment felt by an individual with concern parameter μ, as follows

$$\begin{aligned} r(\mu, s, b) &= h(b - b^*(\mu, s)), \\ h(\cdot) &> 0, \\ h(0) &= 0. \end{aligned}$$

The authors assumed that the level of stigma is aggregate taxpayer resentment as follows

$$s = (N-n)\int_0^{\hat{\mu}} r(\mu, s, b)dG(\mu)$$

Thus, the equilibrium level of stigma must satisfy the equation,

$$s^* = (N-n)\int_0^{\hat{\mu}} r(\mu, s^*, b)dG(\mu)$$

Results in comparative statics is summarised in as the follwoing propostion.

Proposition 3

$$
\begin{aligned}
\frac{ds^*}{db} &> 0, \\
\frac{ds^*}{dy} &< 0, \\
\frac{ds^*}{dn} &\gtreqless 0, \\
\frac{ds^*}{dw} &< 0, \\
\frac{ds^*}{d\gamma} &\gtreqless 0.
\end{aligned}
$$

There exist two essential components in analising comparative statics. The first is the effect on the rich of a change in their income. The second is the effect on the rich of changes in the composition of wefare claimants and incentives to go on welfare.

Stigma and program design

The authors analyse the relationship between stigma and accuracy of targeting. They suppose that the government can identify whether an claimant's type is needy or non-needy in probability τ. Then, the number of poor who claim to take-up welfare is

$$m^\tau := n\left[\gamma + (1-\tau)(1-\gamma)(1-\hat{\theta})\right].$$

Clearly,

$$m^\tau = n\left[\gamma + (1-\tau)(1-\gamma)(1-\hat{\theta})\right] < m = n\left[\gamma + (1-\gamma)(1-\hat{\theta})\right].$$

Under the statistical discrimination view stigma,

$$\frac{ds^*}{d\tau} = \frac{g' \frac{\partial \bar{\theta}_w}{\partial \tau}}{1 - g' \frac{\partial \bar{\theta}_w}{\partial s}} < 0.$$

Next, under the taxpayer resentment view stigma model,

$$\frac{ds^*}{d\tau} = \frac{(N-n) \int_0^{\hat{\mu}} \frac{\partial r}{\partial \tau} dG(\mu)}{1 - (N-n) \int_0^{\hat{\mu}} \frac{\partial r}{\partial s} dG(\mu)} < 0.$$

Proposition 4 *Improved targeting to reduce the number of undeserved claimants reduces welfare stigma under both the the statistical discrimination and the taxpayer resentment views.*

I have to keep in mind that additional costs incurred by raising τ are not considered in this propostion. The authors argue these costs by separating two kinds. The first kind is the cost is driven by additional administrative cost for means testing or inspecting more precisely. The level of stigma could be increased by raising tax under the taxpayer resentment view.

The second is the kind of cost that additional psychic costs generated by treated more differntly by a more targeted program.

Finally, the authors emphasise the possibility that "workfare would reduce welfare stigma in the taxpayer resentment model and would eliminate it in the statistical discrimination model"(P. 181 in Besley and Coate, 1992).

1.2 Survey of empirical researches on welfare stigma

In empirical studies, researchers have tried to estimate the take-up rate of welfare benefit programmes or social benefit programmes (Duclos, 1995; Blank and Ruggles, 1996; Riphahn, 2001; Tachibanaki and Urakawa, 2006; Bargain, Immervoll, and Viitamäki, 2012). For example, Bargain, Immervoll, and Viitamäki (2012) pointed out that non-take-up of welfare has been demonstrated in Finland. In particular, Tachibanaki and Urakawa (2006) suggested take-up rate in Japan is extremely lower than one in other developed countries.

Welfare stigma is regarded to be one of causes of non-take-up welfare. The standard model based on the optimal time allocation problem has not considered stigma. However, Moffitt (1983) analyses the time allocation model includes two stigma costs; fixed stigma and variable stigma with respect to benefit level. Moreover, that paper empirically tested theoretical results using Panel Study of Income Dynamics (PSID). In the consequence, that author suggested that fixed stigma is statistically significant but that variable stigma is not.

In recent years, there emerge some challenging researches related welfare stigma. One of such researches is Bhargava and Manoli (2015). They investigated the causes of non-take-up of welfare, conducted field experiments with the US Internal Revenue Service (IRS) and estimated fractions of welfare take-up using *difference in differences* estimation. They indicated that welfare stigma has a statistically significant impact on welfare take-up rates. Friedrichsen, König, and Schmacker (2018) indicated that the existence of stigma reduces the take-up rate by about 30 percents.

The next section reviews Moffitt (1983) becasue that paper related to chapter 5 in this thesis.

1.2.1 Moffitt (1983)

Most economists use the standard model based on the maximisation problem regarding to two goods; leisure and consumption good. There cannot occur non-take-up welfare in this model. Moffitt (1983) tried modelling "this seemingly irrational rejection of an increase in income" (Moffitt, 1983, p.1023) as resulting from welfare stigma.

Moffitt (1983) modified the standard model to contain welfare stigma. In this paper, the model is developed and estimated for the AFDC programs.

Model

Let us see the model in Moffitt (1983). In this model, stigma is introduced as disutility arising from being on welfare. Moreover, there exist two types of stigma. The first is a flat stigma which arising from the fact of participation, but does not vary with regard to the level of benefit. The second is a variable stigma that varies with regard to the level of welfare.

Suppose that a household's utility without welfare is a function of income,Y:

$$U = U(Y)$$

where $U(\cdot)$ is a monotonic, strictly quasi-concave function in Y. In this model, a flat component of stigma will be said to exist if:

$$\text{Utility} = U(Y + PB) - \phi P$$

where P is a welfare participation variable, $P = \{0, 1\}$, B is the individual's potential welfare benefit. Next, a variable component of stigma will be said to exist if:

$$\text{Utility} = U(Y + \gamma PB)$$

This chapter tested empirically the hypotheses:

$$\phi > 0,$$

$$0 < \gamma < 1.$$

These components of stigma have different implications for a household's decision-making. If only the flat component exists, a household will participate only if

$$U(Y + B) - \phi > U(Y)$$

$$\Leftrightarrow \phi < U(Y + B) - U(Y).$$

That is to say, if the extra utility by participating the welfare benefit is larger than the disutility from stigma generated by doing that, the household will take part in welfare.

The general form of the benefit formula is

$$B = G - tWH - rN,$$

where G is the guarantee income level, W is the household's hourly wage rate, H is hours of work, N is non-wage income, t is the marginal tax rate on earnings.

Next, we consider the case that there is only a variable component of stigma. In this case, the size of γ has important implications for labour supply, for the individual is not different to the mix of earnings and welfare in total income.

We refine the utility function adding labour supply,

$$U(H, Y + \gamma PB) - \phi P,$$

where H is hours of work, then which is called "bad". Budget constraints are

$$Y = WH + N,$$

$$B = G - tWH,$$

$$G = G - rN,$$

where G is the benefit at zero hours of working. Holding P fixed, the first order condition is

$$-\frac{\frac{\partial U}{\partial H}}{\frac{\partial U}{\partial [Y+\gamma PB]}} = (1-\gamma tP)W.$$

If $P = 1$ and $\gamma = 1$, there is no variable stigma, the net wage is $(1-t)W$ as usual condition. However, there is variable stigma ($\gamma < 1$), the net wage is larger than the one in the case of non-variable stigma,

$$(1-\gamma t)W > (1-t)W.$$

It is because that an adding hour of work increases earnings and decreases the welfare benefit in psychological accounting. That is, the marginal utility of earnings relative to the marginal utility of welfare benefit in $\gamma < 1$ is greater than that in $\gamma = 1$. It means one dollar from welfare benefit in the former is relatively more expensive than that in the latter. There can be interesting point that an individual taking-up welfare benefit with "monetary" net wage $(1-t)W$ has a greater incentive to work than an individual not taking-up the benefit with net wage $(1-t)W$. Optimal labour supply function is

$$H = H[(1-\gamma tP)W, N + \gamma\bar{G}],$$

where $(1-\gamma tP)W$ is the psychological real net wage and $N+\gamma G$ is the net non-wage income. Thus, this chapter checked that the presence of variable stigma by testing whether the wage and income elasticities for income of welfare and the one of non-welfare income are different or not.

The choice of participating in the program, P, is determined by which indirect utility is greater under $P = 0$ or $P = 1$. Here, we define $H^*(P)$ as the chosen utility maximising value of H from the direct utility function, Y^*, and B^* as the corresponding amounts of private and welfare income calculated from the budget constraints. Thus, the indirect utility function is

$$V[P, (1 - \gamma t P)W, N + \gamma \bar{G}] = U[H^*(P), Y^* + \gamma B^* P] - \phi P$$

The utility difference P^* is:

$$P^* = V[1, (1 - \gamma t)W, N + \gamma G] - V[0, W, N]$$

If $P^* > 0$, a household participates and not otherwise. To summarise, there are two consumer demand equations in the problem:

$$P = \begin{cases} 1 & \text{if} \quad P^* \geq 0 \\ 0 & \text{if} \quad P^* < 0 \end{cases}$$

$$H = \begin{cases} H[(1 - \gamma t P)W, N + \gamma \bar{G}] & \text{if} \quad P^* \geq 0 \\ H[W, N] & \text{if} \quad P^* < 0 \end{cases}$$

where

$$P^* = V[1, (1 - \gamma t)W, N + \gamma \bar{G}]$$

Figure 1.1 shows a standard labour-leisure diagram with budget constraints ACD off welfare and BC on welfare. In the figure, the utility off welfare is maximised at point E and the hours of work is larger than the eligibility hours level H . If there were no welfare stigma, the individual would choose working hours $H^{*'} < \hat{H}$ and go on welfare as point $E^{'}$ in Figure 1.1. However, if there is stigma, the utility level on welfare is less than the one at point $E^{'}$. The indifference curve corresponds to the dash curve in Figure 1. If this dash indifference curve is below $V[0, W, N]$, the individual would not drop below H

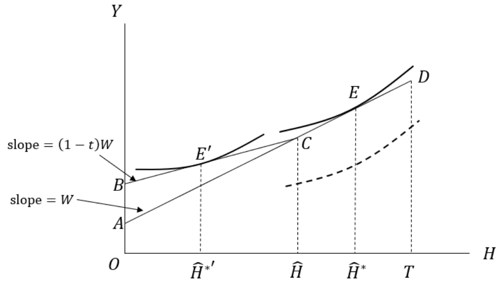

FIGURE 1.1: Working hour-consumption diagram

, while if it were above $V[1, (1-\gamma t), N+\gamma G]$, the household would drop below $\hat{H}$.

Functional form specification

To estimate the system of consumer-demand equations, we need to use specific functional forms for the labour-supply function and the indirect utility function. Consider the linear labour supply function

$$H = \alpha + \beta W_i + \delta N_i$$

where W_i is the net wage rate, $W_i = W$ or $(1-\gamma t)W$, N_i is the net non-wage income, $N_i = N$ or $N + \gamma G$. From the well-known integrability problem, this function can be integrated up to the indirect utility function, V, as given:

$$\begin{aligned} \ln V &= -\ln\left(\beta - \delta\alpha - \delta\beta W_i - \delta^2 N_i\right) - \delta W_i \\ &= -\ln(\beta - \delta H) - \delta W_i \end{aligned}$$

Thus, V is a function of $\ln(\beta - \delta H)$ and δW_i. The former is the compensated substitution effect in the linear model. The latter corresponds to the total income elasticity of labour supply. The indirect utility function is increasing in W_i and N_i, and is strictly convex in W_i so long as$\beta - \delta H > 0$. $\beta - \delta H$ expected to be positive since it is the compensated substitution effect. It will indeed be so if the income effect δ is negative and if the labour-supply curve is upward sloping ($\beta > 0$), which we expect it to be for females in the low-income population. The direct utility function is:

$$\ln U = -\ln(\beta - \delta H) - \frac{[H - \alpha - \delta Y]\delta}{\beta - \delta H}$$

The direct utility function used in this study will be:

$$\ln U = -\ln(\beta - \delta H) - \frac{[H - \alpha - \delta Y]\delta}{\beta - \delta H} - \phi P$$

where γ and ϕ are the stigma parameters as defined in previous section. The indirect utility function is:

$$\ln V_p = -\ln(\beta - \delta H) - \delta W_P - \phi P$$

where

$$W_P = (1 - \gamma t)W, H_P = \alpha + \beta W + \delta(N + \gamma G).$$

We can rewrite the consumer-demand equations as follows:

$$P = \begin{cases} 1 & \text{if} \quad P^* \geq 0 \\ 0 & \text{if} \quad P^* < 0 \end{cases} \tag{1.1}$$

$$H = \begin{cases} H_1 = H[(1 - \gamma tP)W, N + \gamma \bar{G}] & \text{if} \quad P^* \geq 0 \\ H[W, N] & \text{if} \quad P^* < 0 \end{cases} \tag{1.2}$$

where

$$P^* = V[1, (1 - \gamma t)W, N + \gamma \bar{G}] \tag{1.3}$$

If $\phi = 0$, it can be shown that the cutoff value of H_0 must lie above the income eligibility point, implying that all those eligible participate. However, if $\phi > 0$, the cutoff point will be lower than this. For a given H_0, there is a cutoff value of ϕ above which an individual does not participate and below which he does. Therefore, equation 1.3 defines a locus of $H_0 - \phi$ combinations separating participants from nonparticipants.

Data

This subsection explain the data set in that research. Moffitt used the data of the 1976 wave of Michigan Panel Study of Income Dynamics. He restrict sample only the household whom head is female. Observations are 565. Since that model endogenised eligibility, he did not limit sample into eligible household or poor household. The variables used are as follows,

- Employment rate
- Percent of population
- Weekly hours worked (H)
- Net wage
- Non-wage income
- AFDC parameter $(t, G, \overline{G}, \sigma)$
- Years of schooling
- Age
- Family size
- Number of children
- Age of youngest child
- Race

- Local unemployment rate

Since AFDC parameters were unobservable, Moffitt estimated them. He suggested a possibility there existed selection bias.

Econometric specification

Unknown parameters α and ϕ are assumed that

$$\alpha = X\psi + e$$

$$\phi = X\psi + u$$

where $e\tilde{}N(0, \sigma_e^2)$ and $e\tilde{}N(0, \sigma_u^2)$. 1.1, 1.2, 1.3 compose a simultaneous equations system with two endogenous variable, H, P. Since P is binary variable, the system is nonlinear. Therefore, Moffitt used nonlinear maximum likelihood procedures (Heckman, 1977; Lee, 1979). Here, this paper expressed the joint probability of H and P using the likelihood function to remove the selection bias which is simultaneous equation bias caused by endogeneity.

Results

First, Moffitt investigated the case without socio-economic variables. The estimate of flat stigma under this setting is 0.65 and statistically significant, while the one of variable stigma is 1.42 and significant.

Second, this chapter explored the case with these variables. Stigma cost of participation in welfare increase with age, smaller family size, lower level of the unemployment. There exists negative relationship between years of schooling and participation rate.

Simulating T and G in current system, participation rate is 0.36, average working hour is 0.38 in overall sample and 0.10 in participants. Disincentive effect by AFDC is about four hours. Consider the case of reducing tax rate in 0.10. The reduction did not affect participation rate but decreases working hour in 0.3 hours. Suppose the case of raising the level of guarantee level to 65% of poverty line. Then, participation rate increased in 0.11 caused by the eligible probability increased to 0.92. Moreover, since income elasticity is statistically significant, working hours of participants reduces in four hours weekly, and those of overall reduces in two hours. An increase in real wage rate did not affect participation rate but increased working hours of participants. An decrease in unemployment rate had strictly positive effect on working hours. Increases in education years and age had positive effect as same as a reduce in unemployment rate. Therefore, change in parameters which are not AFDC parameters had larger impact on labour supply than that in AFDC parameters.

Chapter 2

Models of endogenous welfare stigma (I): Statistical discrimination view

Abstract

This study extends the statistical discrimination stigma model to analyse not only welfare fraud but also non-take-up of welfare benefits. Specifically, I extend the statistical stigma model to endogenise the decision-making process for the 'needy type'. The comparative static analysis indicates that an increase in the benefit level can make non-take-up of welfare benefits more serious. I demonstrate this result through a simple equation, linking two types of elasticity for needy and non-needy types.

2.1 Introduction

Empirical and experimental studies have considered welfare stigma to be crucial in understanding the impact of welfare programmes with regard to the goal of alleviating poverty (Moffitt, 1983; Friedrichsen, König, and Schmacker, 2018).

Besley and Coate (1992) pioneering research analysed situations wherein stigmas were endogenised. 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 United Kingdom was approximately 80 per cent (Duclos, 1995), approximately 60–67 per cent in the United States (Blank and Ruggles, 1996), approximately 37 per cent in Germany Riphahn2001 and 16.3–19.7 per cent in Japan (Tachibanaki and Urakawa, 2006).

Blumkin, Margalioth, and Sadka (2015) analysed welfare stigma as a policy tool, which was used to restrain welfare fraud. Thus, non-take-up welfare did not manifest in their model.

This study extends Besley and Coate (1992)'s model to explain the occurrence of non-take-up of welfare benefits. Unlike Besley and Coate (1992)'s, I endogenise decision-making for non-workers. The comparative static analysis indicates that an increase in the benefit level makes non-take-up of welfare benefits a serious concern.

2.2 Model

In this section, I develop a model that analyses interactions among welfare fraud, non-take-up welfare, and welfare stigma.

2.2.1 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 hope so. I 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, I 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 $\phi_i(0,1)$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. I assume the following properties, for $\forall I$, $\forall z_i$, $i = 1,2$,

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

where I is the level of income and $I \in w, b^1$. 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:

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

[1]For simplicity, I assume the price of consumption good is 1

Here θ is disutility of labour, and w is work income. I 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.

2.2.2 The critical level of sensitivity to stigma

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

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

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:

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

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

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

2.2.3 Formulation of the stigma cost function

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

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

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

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

This formulation is inspired by the statistical stigma in Besley and Coate (1992) and Blumkin, Margalioth, and Sadka (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 commited.

Stigma cost is a function of capability. While Besley and Coate (1992) assumed that stigma cost was the same for all recipients, I differentiate stigma cost by the capabilities of type 1 and 2. Even though, I do not assume the sign of $\partial s\left(\Pi, z_i\right)/\partial z_i$, in each case I explain whether $\partial s\left(\Pi, z_i\right)/\partial z_i > 0$ or not. When $\partial s\left(\Pi, z_i\right)/\partial z_i > 0$, stigma costs for non-needy types are higher than those for needy types. I 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, I obatain the following:

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

In the next section, I show this model's equilibrium.

2.3 Equilibrium

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

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

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

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

Here,

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

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

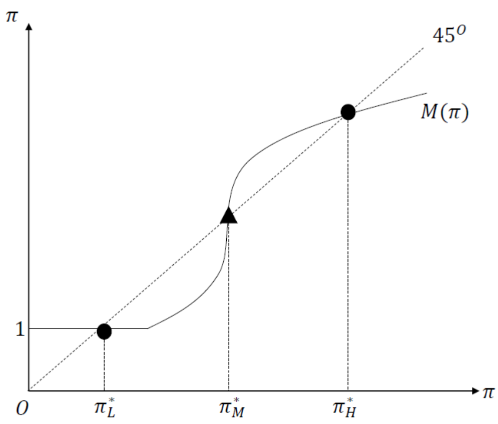

FIGURE 2.1: Example of multiple equilibria

$\hat{u}(b, z_i)$ is the incremental material utility when taking-up welfare. $M(\pi)$ is a mapping from π to itself. By differentiation, I 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, I 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, I rewrite this as given:

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

Equation 2.3 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

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

Figure 2.1 shows the case of multiple equilibria.

2.4 Comparative statics

In this section, I conduct comparative statics. I am particularly interested in how a change in benefit level to an equilibrium.

I 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 summarised in the following proposition.

Proposition 5

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

$$\frac{dp^*}{db} < 0 \; \textit{if} \; \frac{\eta_b(z_1)}{\eta_b(z_2)} < \frac{\varepsilon_{\pi^*}(z_1)}{1+\varepsilon_{\pi^*}(z_2)}, \; \frac{dp^*}{db} \geq 0 \; \textit{otherwise}\,.$$

Proof. Equilibrium equations are as follows:

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

By logarithmic transformation, we obtain the following:

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

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

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

$$\Longleftrightarrow$$

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

$$\Longleftrightarrow$$

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

A matrix representation is given below:

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

By Cramer's rule.

$$\frac{dp/p}{db/b} = \frac{\begin{vmatrix} \eta_b\left(z_1\right) & 0 & -\varepsilon_\pi\left(z_1\right) \\ \eta_b\left(z_2\right) & 1 & -\varepsilon_\pi\left(z_2\right) \\ 0 & -1 & -1 \end{vmatrix}}{\begin{vmatrix} 1 & 0 & -\varepsilon_\pi\left(z_1\right) \\ 0 & 1 & -\varepsilon_\pi\left(z_2\right) \\ 1 & -1 & -1 \end{vmatrix}} = \frac{-\eta_b\left(z_1\right)\left[1 + \varepsilon_\pi\left(z_2\right)\right] + \eta_b\left(z_2\right)\varepsilon_\pi\left(z_1\right)}{\varepsilon_\pi\left(z_1\right) - \left[1 + \varepsilon_\pi\left(z_2\right)\right]}.$$

Since the stability condition is $\varepsilon_{\pi^*}(z_1) - \varepsilon_{\pi^*}(z_2) < 1$, the denominator is negative. Therefore, we get the following:

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

■

When the elasticity of non-needy people of incremental material utility to his/her benefit level is relatively high or the elasticity of needy people of stigma cost to the ratio of p to q, π is relatively high, an increase in the benefit level is likely to reduce take-up of welfare benefits in needy types.

2.5 Conclusion

This chapter has explored and extended Besley and Coate (1992)'s statistical discrimination model. The model can explains the needy type's decision not to take-up welfare. In a comparative static analysis, I find that higher benefit levels can reduce the number of recipients in the equilibrium. Two elasticities in each type are also shown and are important toward understanding the results. One is the elasticity of stigma cost to the ratio of deserving beneficiaries to undeserved beneficiaries. The other is the elasticityof material utility to benefit levels.

Chapter 3

Models of endogenous welfare stigma (II): Relative income resentment view

Abstract

The third chapter studies the stigma model of relative income. The model is similar to the taxpayer resentment model (Besley and Coate, 1992) in the point that non-recipients, i.e., workers, taxpayers or capitalists, get anger and such feelings creates stigma. The model differs from Besley and Coate (1992)'s taxpayer resentment model in that the stigmatisee can become the stigmatiser. Moreover, I use relative income mechanism to formulate stigma which is increasing function with the ratio of benefit level to each working wage. As a result, multiple equilibria can occur, high stigma and low stigma, contrary to the Besley and Coate (1992)'s model. In the comparative static analysis, the result indicates that the possibility that the number of recipients can decline when a negative macroeconomic shock occurs against an intuition.

3.1 Introduction

Although the general belief is that poor people are more supportive of generous redistribution (Meltzer and Richard, 1981), it does not appear to be consistent with the actual situation. In the 2016 United States presidential election, many blue-collar workers supported Donald John Trump, who led the effort to repeal and replace the Affordable Care Act (also known as 'Obamacare').

Shayo (2009) attempts to shed the light on why blue-collar worker in the United States are less supportive of redistribution compared with their German counterparts. He applies the social identity model to a standard framework in the political economy model to prove the existence of multiple equilibria, wherein one equilibrium suggests relatively high levels of redistribution and class identification among the poor and the other correlates with comparatively low levels of redistribution and national identification. This paper tries to explain such phenomena by another mechanism, i.e., stigma from relative income resentment.

The model is similar to the taxpayer resentment model (Besley and Coate, 1992) in the point that non-recipients (workers, taxpayers or capitalists) get anger and such a feeling creates stigma. However, the model differs from Besley and Coate (1992)'s taxpayer resentment model in that the stigmatisee can become the stigmatiser. Moreover, I use relative income mechanism to formulate stigma which is increasing function with the ratio of benefit level to each working wage. In the comparative static analysis, the result indicates that the possibility that the number of recipients can decline when a negative macroeconomic shock occurs against an intuition.

The paper is structured as follows. The next section sets up the model, while the third section defines the equilibrium. The fourth section investigates the comparative statics, and the final section concludes this paper.

3.2 Basic setting of the model

We consider the mass of population to be normalised to one and denote an index of individuals $i \in [0, 1]$. Each individual can either work or take-up welfare. If an individual chooses to work, he/she earns a working income $w(i, y)$, and y is a parameter for the macroeconomic environment. We assume that

$$\frac{\partial w(i,y)}{\partial i} > 0,$$
$$\frac{\partial w(i,y)}{\partial y} > 0.$$

The setting of utility is

$$\begin{cases} u(w(i,y)) - \theta & \text{if working} \\ u(b) - \phi s & \text{if taking-up welfare,} \end{cases}$$

Here b is the benefit level, θ is disutility of labour, s is the level of stigma, and ϕ is the sensitivity to stigma.

The stigma experienced by recipients is generated by workers' resentment. If an individual i chooses to work, i's resentment is

$$\sigma\left(\frac{b}{w(i,y)}\right),$$

where $\sigma^{'}(\cdot) > 0$. This reflects that worker's resentment increases with the ratio of benefit level to wage rate. The stigma level is assumed to be aggregate of resentment as follows:

$$s = \int_{\hat{\imath}}^{1} \sigma\left(\frac{b}{w(i,y)}\right) di.$$

Here, $\hat{\imath}$ is the index of workers whose wage income is at the minimum level.

3.3 Equilibrium

We suppose that for individuals who take-up welfare $i \in [0, \bar{\imath}]$, an equilibrium level of threshold wage income, $\hat{\imath}^*$, is equal to the solution obtained from the following equations.

We suppose that individuals $i \in [0, \bar{\imath}]$ take-up welfare. An equilibrium level of threshold wage income, $\hat{\imath}^*$, is equal to the solution obtained from the following equations.

$$\begin{cases} u(b) - \phi s = u(w(\hat{\imath}, y)) - \theta, \\ s = \int_{\bar{\imath}}^{1} \sigma(\frac{b}{w(i,y)}) di, \\ \bar{\imath} = \hat{\imath}. \end{cases}$$

By totally differentiating,

$$\frac{d\hat{\imath}}{d\bar{\imath}} = \frac{\phi\sigma(b/w(\bar{\imath}, y))}{u'(w(\hat{\imath}, y))\frac{\partial w(\hat{\imath},y)}{\partial \hat{\imath}}},$$

the stable condition is as follows:

$$\frac{\phi\sigma\left(b/w\left(\hat{\imath}^*, y\right)\right)}{u'\left(w\left(\hat{\imath}^*, y\right)\right)\frac{\partial w(\hat{\imath}^*,y)}{\partial \hat{\imath}}} < 1.$$

3.4 Comparative statics

In this section, I conduct a comparative static analysis. First, I examine how changes in benefit levels impact equilibrium. In equilibrium, the following equation is satisfied:

$$u(b) - \phi \int_{\hat{\imath}^*}^{1} \sigma(b/w(i, y)) di = u\left(w\left(\hat{\imath}^*, y\right)\right) - \theta$$

Totally differentiating and assuming $dy = d\theta = d\phi = 0$,

$$\left[u'(b) - \phi \int_{\tilde{i}^*}^{1} \sigma' \left(\frac{b}{w(i,y)}\right) di\right] db + \phi\sigma \left(\frac{b}{w\left(\hat{\imath}^*, y\right)}\right) di^* = u'\left(w\left(\hat{\imath}^*, y\right)\right) \frac{\partial w\left(\hat{\imath}^*, y\right)}{\partial \hat{\imath}} d\hat{\imath}^*$$

$$\frac{di^*}{db} = \frac{u'(b) - \phi \int_{i^*}^{1} \sigma' \left(\frac{b}{w(i,y)}\right) di}{u'\left(w\left(\hat{\imath}^*, y\right)\right) \frac{\partial w\left(\hat{\imath}^*, y\right)}{\partial i} - \phi\sigma\left(b/w\left(\hat{\imath}^*.y\right)\right)} \tag{3.1}$$

By a stable condition, the denominator to the right-hand side of equation 3.1 is positive.

Proposition 6 *Under the relative income stigma, an increase in benefit levels can either increase or decrease the number of recipients.*

$$\frac{di^*}{db} = \frac{u'(b) - \phi \int_{i^*}^{1} \sigma' \left(\frac{b}{w(i,y)}\right) di}{u'\left(w\left(\hat{\imath}^*, y\right)\right) \frac{\partial w\left(\hat{\imath}^*, y\right)}{\partial i} - \phi\sigma\left(b/w\left(\hat{\imath}^*, y\right)\right)},$$

$$\operatorname{sign} \frac{d\hat{\imath}^*}{db} = \operatorname{sign} \left[u'(b) - \phi \int_{i^*}^{1} \sigma' \left(\frac{b}{w(i,y)}\right) di\right]. \tag{3.2}$$

The first term in the right-hand side of equation 3.2 is a positive direct effect: the number of recipients increases with a rise in benefit levels. The second term is a negative indirect effect: the number of recipients decreases with an increase in stigma due to rising resentment of workers. When ϕ is sufficiently high, the indirect effect is likely to dominate the direct effect. That is, when sensitivity to stigma is high, contrary to intuition, the equilibrium level of recipients decreases with benefit levels.

Second, we investigate how changes in the macroeconomic environment impact the equilibrium levels of the threshold. Totally differentiating and assuming $db = d\theta = d\phi = 0$,

$$\left[\phi\int_{i^*}^{1}\sigma'(b/w(i,y))\frac{b}{w(i,y)^2}\frac{\partial w(i,y)}{\partial y}di\right]dy+\phi\sigma\left(b/w\left(\hat{\imath}^*,y\right)\right)d\hat{\imath}^*$$
$$=u'\left(w\left(\hat{\imath}^*,y\right)\right)\left[\frac{\partial w\left(\hat{\imath}^*,y\right)}{\partial y}dy+\frac{\partial w\left(\hat{\imath}^*,y\right)}{\partial\hat{\imath}}d\hat{\imath}^*\right]$$

$$\frac{d\hat{\imath}^*}{dy}=\frac{\phi\int_{\hat{\imath}^*}^{1}\sigma'\left(b/w\left(\hat{\imath}^*,y\right)\right)\frac{b}{w(\hat{\imath}^*,y)^2}\frac{\partial w(\hat{\imath}^*,a)}{\partial y}di-u'\left(w\left(\hat{\imath}^*,y\right)\right)\frac{\partial w(\hat{\imath}^*,y)}{\partial y}}{u'\left(w\left(\hat{\imath}^*,y\right)\right)\frac{\partial w\left(\hat{\imath}^*,y\right)}{\partial\hat{\imath}}-\phi\sigma\left(b/w\left(\hat{\imath}^*,y\right)\right)}$$

Proposition 7 *In the relative income stigma, macroeconomic shock may either increase or decrease the number of recipients.*

$$\frac{d\hat{\imath}^*}{dy}=\frac{\phi\int_{\hat{\imath}^*}^{1}\sigma'\left(b/w\left(\hat{\imath}^*,y\right)\right)\frac{b}{w(\hat{\imath}^*,y)^2}\frac{\partial w(\hat{\imath}^*,a)}{\partial y}di-u'\left(w\left(\hat{\imath}^*,y\right)\right)\frac{\partial w(\hat{\imath}^*,y)}{\partial y}}{u'\left(w\left(\hat{\imath}^*,y\right)\right)\frac{\partial w\left(\hat{\imath}^*,y\right)}{\partial\hat{\imath}}-\phi\sigma\left(b/w\left(\hat{\imath}^*,y\right)\right)},$$

$$\operatorname{sign}\frac{d\hat{\imath}^*}{db}=\operatorname{sign}\left[\phi\int_{\hat{\imath}^*}^{1}\sigma'\left(b/w\left(\hat{\imath}^*,y\right)\right)\frac{b}{w\left(\hat{\imath}^*,y\right)^2}\frac{\partial w\left(\hat{\imath}^*,a\right)}{\partial y}di-u'\left(w\left(\hat{\imath}^*,y\right)\right)\frac{\partial w\left(\hat{\imath}^*,y\right)}{\partial y}\right] \tag{3.3}$$

In the right hand side of equation3.3, the second term corresponds to the negative direct effect that recipients reduced by positive macroeconomic shock. The first term represents the positive indirect effect on recipients, which increases in case of decrease in stigma from reduced resentment due to improvement in macroeconomic shock. In other words, when negative macroeconomic shock occurs, contrary to intuition, the equilibrium number of recipients may decrease.

3.5 Conclusion

The third chapter studied the stigma model of relative income. The model was similar to the taxpayer resentment model (Besley and Coate, 1992) in the point that non-recipients (workers, taxpayers or capitalists). The model differed from Besley and Coate (1992)'s taxpayer resentment model in that the stigmatisee can become the stigmatiser. Moreover, we used relative income mechanism to formulate stigma which was increasing function with average of the ratio of benefit level to each working wage. As a result, multiple equilibria could occur, high stigma and low stigma, contrary to the proportion of Besley and Coate (1992)'s model. It was because there exists feedback effect. In the comparative static analysis, the result indicated that the possibility that the number of recipients declined in case of a negative macroeconomic shock since there existed the negative indirect effect on recipients, which decreased in case of increase in stigma from increased resentment due to a shift of working income distribution.

Chapter 4

Aspiration enhancing role of stigma and social mobility

Abstract

This chapter researches the aspiration enhancing role of the stigma created by low income and poverty. I discuss the possibility that this stigma gives children an incentive to seek an escape from poverty, and I investigate the cultural mechanism of transmitting this stigma inter-generation. In some cases, increased income inequality immobilises family cultural capital, and social mobility is constrained.

In this paper, I shed light on the aspiration enhancing role of stigma and social mobility, presenting a model in which the sensitivity to stigma is endogenous. Specifically, humans endogenise forming preferences using the cultural transmission model. With this, I focus on the relationship between poverty and culture. The idea was inspired by Oscar Lewis, an American anthropologist who discussed the culture of poverty (Lewis, 1966). Many economists and sociologists have also researched social mobility. Smeeding, Erikson, and Jäntti (2011) indicated that there is a positive correlation between parents' and children's income in the United States and United Kingdom. In theoretical works, Piketty (1995) and Benabou and Tirole (2006) analysed models in which social mobility, belief, and redistributive politics were determined endogenously. This paper follows Besley (2016), who analysed a model of aspiration enhancing role in which parents' income level was used as a reference point for children's preferences. In that model, children were socialised to be aspiration-enhanced or not. They considered situations for which there are aspiration enhancing incentives for both stigma-sensitive and stigma-insensitive children. Both types can receive incentives to socialise cultural traits. Therefore, I use the framework of cultural transmission. The framework was developed by Bisin and Verdier (2000) and Bisin and Verdier (2001), who modeled parents' cultural socialisation of children. Their formulation of imperfect empathy is an important concept in which parents assess their children's utility through their own preferences. I partially extend this idea to parents empathizing with their children. I denote γ as the parameter of the degree of empathy from parents to children, giving us interesting suggestions and implications. The structure of this paper as follows. Section 2 explains the model. In section 3, I model the socialisation mechanism. In section 4, I evaluate the steady state. Section 5 presents a comparative static analysis. Finally, I conclude this paper in Section 6.

4.1 Model

Consider an overlapping-generation economy in which each generation lives for two periods. There are two generations: young and adult. The young generation makes effort and they obtain income when adult. Let e=0,1 be effort choice. I assume that a unit of labor is inelastically supplied when a youth becomes an adult and earns income y_H or y_L, $y_H > y_L > 0$. The probability of child earning a high income as an adult is $\theta(e)$, $\theta(\cdot)$. Thus, $\theta(1) > \theta(0) > 0$,$\theta_H := \theta(1)$,$\theta_L := \theta(0)$.

There are two types of people, stigma sensitive and insensitive, denoted as type 1 and type 0, respectively. Type 1's proportion in the population is q_t during period t. The utility of a child with $\tau \in 1, 0$is given as,

$$U = \delta \left\{ \theta(e) y_H + (1 - \theta(e)) (y_L + b) \right\} - \sigma(\tau) (1 - \theta(e)) S - eK$$

where $\delta \in [0, 1]$ is the index myopia of a child, K is the effort cost, S is the stigma, $\sigma(\tau)$ is the indicator variable, $\sigma(1) = 1$, $\sigma(0) = 0$. $\delta < 1$ when young and $\delta = 1$ when an adult. The utility of children has trait $\tau \in 1, 0$ $\tilde{u_\tau}$,

$$\tilde{u_\tau} = \delta \left\{ \theta(e) y_H + (1 - \theta(e)) (y_L + b) \right\} - \sigma(\tau) (1 - \theta(e)) S - eK$$

and adults with $\tau \in 1, 0$, u_τ,

$$u_\tau = \left\{ \theta(e) y_H + (1 - \theta(e)) (y_L + b) \right\} - \sigma(\tau_i) (1 - \theta(e)) S - eK$$

Thus, I assume that

$$(\theta_H - \theta_L) \delta (y_H - y_L) < K$$

Thereore, the insensitive type's optimal effort choice is $e_0 = 0$. Here, let S_{min} denote the value of S such that

$$(\theta_H - \theta_L) [\delta (y_H - y_L) + S_{\min}] = K.$$

Then,

$$e_1 = \begin{cases} 0 & \text{if } S < S_{\min} \\ 1 & \text{if } S \geq S_{\min} \end{cases}$$

I assume that

$$(\theta_H - \theta_L) (y_H - y_L) > K,$$

Thus, the effort is efficient within the economy. I suppose stigma that the stigma is the the following function,

$$S = S\left(q_t\right)$$

I assume the property, $S() > 0$. This property reflects that stigma-sensitive type is larger, stigma from economic failure is higher.

4.2 Socialisation

In this section, consider cultural transmission as value formation. Following Bisin and Verdier (2000) and Bisin and Verdier (2001)'s cultural transmission mechanism, the probability, P^{ij} that a child from a family with trait $\tau = i \in 1,0$ will be socialised to trait $\tau \in 1,0$ is shown as

$$P^{11} = d_{1t}+(1-d_{1t})q_t P^{10} = (1-d_{1t})(1-q_t) P^{00} = d_{0t}+(1-d_{0t})(1-q_t) P^{01} = (1-d_{0t})q_t$$

where, $d_{\tau t}$ is the direct socialisation level of a parent with $\tau \in 0,1$. Parents with trait i maximise

$$\max_{d_i} P^{ii}(d_i)V^{ii} + P^{ij}(d_i)V^{ij} - c(d_i),$$

where

$$V^{ij} = \gamma u_j(e_j) + (1-\gamma)u_i(e_j)e_j = argmax_e u_j(e)$$

This is an extended version of imperfect empathy, as provided by Bisin and Verdier (2000) and Bisin and Verdier (2001). Their model was equal to the $\gamma = 0$ case of our model. Thus, the parameter reflects the index of the degree of empathy from parents to children.

$$
\begin{aligned}
u_s\left(e_s^*\right) &= \left\{\theta_H y_H + \left(1-\theta_H\right)\left(y_L+b\right)\right\} - \left(1-\theta_H\right)S - K \\
u_s\left(e_n^*\right) &= \left\{\theta_L y_H + \left(1-\theta_L\right)\left(y_L+b\right)\right\} - \left(1-\theta_L\right)S \\
u_n\left(e_n^*\right) &= \left\{\theta_L y_H + \left(1-\theta_L\right)\left(y_L+b\right)\right\} \\
u_n\left(e_s^*\right) &= \left\{\theta_H y_H + \left(1-\theta_H\right)\left(y_L+b\right)\right\} - K
\end{aligned}
$$

The costs of socialisation are formulated as follows. For parents with trait $\tau \in 0, 1$,

$$c(d_\tau) = \frac{d_\tau^2}{2\phi}.$$

4.2.1 The case of $S \geq S_{\min}$

For parents with trait $\tau = 1$

$$
\begin{aligned}
V^{11} &= \theta_H y_H + \left(1-\theta_H\right)y_L - K - \left(1-\theta_H\right)S \\
V^{10} &= \theta_L y_H + \left(1-\theta_L\right)y_L - \left(1-\theta_L\right)(1-\gamma)S
\end{aligned}
$$

$$
\begin{aligned}
\Delta V^s &= V^{ss} - V^{sn} \\
&= \left[\left(\theta_H-\theta_L\right)\left(y_H-y_L\right) - K\right] - \left[\left(1-\theta_H\right) - \left(1-\theta_L\right)\left(1-\gamma\right)\right]S \\
&= \overline{y} + \left[\left(\theta_H-\theta_L\right) - \left(1-\theta_L\right)\gamma\right]S
\end{aligned}
$$

where $y := (\theta_H - \theta_L)(y_H - y_L) - K$. When

$$(\theta_H - \theta_L) - (1-\theta_L)\gamma < 0,$$

$\leftrightarrow \gamma \geq \frac{\theta_H - \theta_L}{1-\theta_L}$,

ΔV^1 is decreasing in S and

$$d_{1t}^* = \begin{cases} \phi\left(1-q_t\right)\Delta V^1 & \text{if } S < \hat{S}_1 \\ 0 & \text{if } S \geq \hat{S}_1 \end{cases}$$

where

$$\hat{S}_1 := \frac{\overline{y}}{\left(1-\theta_L\right)\gamma - \left(\theta_H - \theta_L\right)}$$

When

$$\gamma < \frac{\theta_H - \theta_L}{1-\theta_L},$$

ΔV^1 is always positive and increasing in S, then,

$$d_1^* = \phi\left(1-q_t\right)\Delta V^1$$

For parents with trait $\tau = 0$

$$V^{00} = \theta_L y_H + \left(1-\theta_L\right) y_L$$
$$V^{01} = \theta_H y_H + \left(1-\theta_H\right) y_L - K - \left(1-\theta_H\right)\gamma S$$
$$\Delta V^0 = V^{00} - V^{01} = \left(1-\theta_H\right)\gamma S - \overline{y}$$

Hence, ΔV^0 is increasing in S.

$$d_n^* = \begin{cases} 0 & \text{if } S < \hat{S}_0 \\ \phi q_t \Delta V^0 & \text{if } S \geq \hat{S}_0 \end{cases}$$

where

$$\hat{S}_0 := \frac{\overline{y}}{\left(1-\theta_H\right)\gamma}$$

4.2.2 The case of $S < S_{\min}$

For parents with trait $\tau = 1$

$$V^{ss} = \theta_L y_H + (1-\theta_L)\, y_L - (1-\theta_L)\, S$$
$$V^{sn} = \theta_L y_H + (1-\theta_L)\, y_L - (1-\theta_L)\,(1-\gamma)S$$
$$\Delta V^s = V^{ss} - V^{sn} = -\,(1-\theta_L)\,\gamma S$$

Therefore,

$$d_1^* = 0$$

For parents with trait $= 0$

$$V^{00} = \theta_L y_H + (1-\theta_L)\, y_L$$
$$V^{01} = \theta_L y_H + (1-\theta_L)\, y_L - (1-\theta_L)\,\gamma S$$
$$\Delta V^0 = V^{00} - V^{01} = (1-\theta_L)\,\gamma S > 0$$

Hence, ΔV^0 is increasing in S.

$$d_0^* = \phi q_t\,(1-\theta_L)\,\gamma S$$

4.2.3 Optimal direct socialisation

I assume that the minimum level of of stigma is not so low, $S(0) > S_{min}$. Each optimal level of direct socialisation is

$$d_0\,(q_t) = \begin{cases} 0 & \text{if } S < S_1 \\ \phi q_t\,[(1-\theta_H)\,\gamma S\,(q_t) - \overline{y}] & \text{if } S \geq S_1 \end{cases}$$

$$d_1(q_t) = \begin{cases} \phi(1-q_t)\left[\overline{y} + \left[(\theta_H - \theta_L) - (1-\theta_L)\gamma\right] S(q_t)\right] & \text{if} \quad \gamma < \frac{\theta_H - \theta_L}{1-\theta_L} \\ & \text{or} \quad S < \hat{S}_1 \wedge \gamma \geq \frac{\theta_H - \theta_L}{1-\theta_L}, \\ 0 & \text{if} \quad S \geq \hat{S}_1 \wedge \gamma \geq \frac{\theta_H - \theta_L}{1-\theta_L}. \end{cases}$$

I define $\hat{q}_0$, $\hat{q}_1$, as follows

$$\hat{S}_0(q)^{-1} := \widehat{q}_0$$

$$\hat{S}_1(q)^{-1} := \widehat{q}_1$$

Since

$$\frac{\partial d_0}{\partial q_t} = \begin{cases} 0 & \text{if } S < \hat{S}_0, \\ \phi\left[(1-\theta_H)\gamma S(q_t) - \overline{y}\right] + \phi q_t (1-\theta_H)\gamma S'(q_t) & \text{if } S \geq \hat{S}_0, \end{cases}$$

d_{0t} is increasing and convex with q_t, as shown in Figure 4.1.

In the case of $\gamma < \frac{\theta_H - \theta_L}{1-\theta_L}$, a change in optimal level of direct socialisation for sensitive type with respect to the proportion of its type to population is given as follows

$$\begin{aligned} \frac{\partial d_1}{\partial q_t} &= -\phi\left[\overline{y} + \left[(\theta_H - \theta_L) - (1-\theta_L)\gamma\right] S(q_t)\right] \\ &\quad + \phi(1-q_t)\left[(\theta_H - \theta_L) - (1-\theta_L)\gamma\right] S'(q_t) \end{aligned}$$

In the right hand side, the first term is negative, and second term is positive. The former corresponds to the complementary effects of increasing vertical socialisation. The latter corresponds to the effects of increasing incentives for direct socialisation with an increased level of stigma cost. Because $d_1(0) = 0$, I can derive an optimal level of direct socialisation for the sensitive type with low empathy, as shown in 4.2.

In the case of $\gamma \geq \frac{\theta_H - \theta_L}{1-\theta_L}$ is given

$$\frac{\partial d_1^*}{\partial q_t} = \begin{cases} -\phi\left[\overline{y} + \left[(\theta_H - \theta_L) - (1-\theta_L)\gamma\right] S(q_t)\right] & \text{if } S < \hat{S}_1 \\ +\phi(1-q_t)\left[(\theta_H - \theta_L) - (1-\theta_L)\gamma\right] S'(q_t) & \text{if } S < \hat{S}_1 \\ 0 & \text{if } S \geq S_1 \end{cases}$$

Therefore, d_1 is decreasing and concave with q_t as shown in Figure 4.3.

4.3 Cultural transmission

The fraction of offspring who will become stigma-sensitive during period $t+1$ as follows:

$$\begin{aligned} q_{t+1} &= q_t P^{11} + (1-q_t) P^{10} \\ &= q_t(1-q_t)(d_{1t} - d_{0t}) + q_t \end{aligned}$$

The difference equation is

$$\Delta q_t = q_{t+1} - q_t = q_t(1-q_t)(d_{1t} - d_{0t})$$

In next two sections, I consider the steady states divided into two cases: high and low empathy case.

4.3.1 Case of high empathy

Because

$$\begin{aligned} &(1-\theta_H)\gamma - \left[(1-\theta_L)\gamma - (\theta_H - \theta_L)\right] \\ &= (\theta_H - \theta_L)(1-\gamma) > 0 \end{aligned}$$

I have

$$\hat{S}_0 = \frac{(\theta_H - \theta_L)(y_H - y_L) - K}{(1-\theta_H)\gamma} < \frac{(\theta_H - \theta_L)(y_H - y_L) - K}{(1-\theta_L)\gamma - (\theta_H - \theta_L)} = S_1.$$

When $S < S_0$,

$$d_{1t}^{*} - d_{0t}^{*} = d_{1t}^{*} = \phi\left(1 - q_t\right)\left[\overline{y} + \left[\left(\theta_H - \theta_L\right) - \left(1 - \theta_L\right)\gamma\right] S\left(q_t\right)\right] > 0$$

When $\hat{S}_0 \leq S < \hat{S}_1$,

$$\begin{aligned} d_{1t}^{*} - d_{0t}^{*} = {} & \phi\left(1 - q_t\right)\left[\overline{y} + \left[\left(\theta_H - \theta_L\right) - \left(1 - \theta_L\right)\gamma\right] S\left(q_t\right)\right] \\ & - \phi q_t\left[\left(1 - \theta_H\right)\gamma S - \overline{y}\right] \end{aligned}$$

is monotonically decreasing in q_t.

When $S > \hat{S}_1$,

$$d_{1t}^{*} - d_{0t}^{*} = -d_{0t}^{*} = -\phi q_t\left[\left(1 - \theta_H\right)\gamma S - \overline{y}\right] < 0$$

The steady state q^{*} is characterised as

$$\phi\left(1 - q^{*}\right)\left[\overline{y} + \left[\left(\theta_H - \theta_L\right) - \left(1 - \theta_L\right)\gamma\right] S\left(q^{*}\right)\right] = \phi q^{*}\left[\left(1 - \theta_H\right)\gamma S - \overline{y}\right]$$
$$\tfrac{q^{*}}{1-q^{*}} = \tfrac{\overline{y}+\left[\left(\theta_H-\theta_L\right)-\left(1-\theta_L\right)\gamma\right]S\left(q^{*}\right)}{\left(1-\theta_H\right)\gamma S-\overline{y}}$$

Figure 4.5 shows a steady state in the case of high empathy.

4.3.2 Case of low empathy

When $S < \hat{S}_0$

$$d_{1t} - d_{0t} = d_{1t} = \phi\left(1 - q_t\right)\left[\overline{y} + \left[\left(\theta_H - \theta_L\right) - \left(1 - \theta_L\right)\gamma\right] S\left(q_t\right)\right] > 0$$

When $S \geq \hat{S}_0$,

$$
\begin{aligned}
d_{1t} - d_{0t} = {} & \phi \left(1 - q_t\right) \left[\overline{y} + \left[\left(\theta_H - \theta_L\right) - | \left(1 - \theta_L\right) \gamma\right] S\left(q_t\right)\right] \\
& - \phi q_t \left[\left(1 - \theta_H\right) \gamma S - \overline{y}\right]
\end{aligned}
$$

The steady state q^* is characterised as

$$
\phi \left(1 - q^*\right) \left[\overline{y} + \left[\left(\theta_H - \theta_L\right) - \left(1 - \theta_L\right) \gamma\right] S\left(q^*\right)\right] = \phi q^* \left[\left(1 - \theta_H\right) \gamma S - \overline{y}\right],
$$

$$
\frac{q^*}{1 - q^*} = \frac{\overline{y} + \left[\left(\theta_H - \theta_L\right) - \left(1 - \theta_L\right) \gamma\right] S\left(q^*\right)}{\left(1 - \theta_H\right) \gamma S - \overline{y}}.
$$

Figure 4.4 shows a steady state in the case of high empathy.

4.4 Comparative statics

This section presents a comparative static analysis with respect to the equilibrium state (q^*, d_0^*, d_1^*). First, I derive the stable equilibrium condition.

$$
\frac{\partial d_{1t}^*}{\partial q_t} - \frac{\partial d_{0t}^*}{\partial q_t} < 0.
$$

$$
\begin{aligned}
\frac{\partial d_{1t}^*}{\partial q_t} - \frac{\partial d_{0t}^*}{\partial q_t} = {} & \phi \left(1 - q^*\right) \left[\overline{y} + \left[\left(\theta_H - \theta_L\right) - \left(1 - \theta_L\right) \gamma\right] S\left(q^*\right)\right] \\
& - \phi q^* \left[\left(1 - \theta_H\right) \gamma S - \overline{y}\right] \\
= {} & - \phi \left[\overline{y} + \left[\left(\theta_H - \theta_L\right) - \left(1 - \theta_L\right) \gamma\right] S\left(q^*\right)\right. \\
& + \phi \left[1 - q^*\right) \left[\left(\theta_H - \theta_L\right) - \left(1 - \theta_L\right) \gamma\right] S\left(q^*\right)\left.\right] \\
& - \phi q^* \left(1 - \theta_H\right) \gamma S'\left(q^*\right) \\
= {} & \phi \left[\left(1 - q^*\right) \left[\left(\theta_H - \theta_L\right) - \left(1 - \theta_L\right) \gamma\right] - q^* \left(1 - \theta_H\right) \gamma\right] S'\left(q^*\right) + 2 \phi \overline{y}
\end{aligned}
$$

$$
S'\left(q^*\right) < - \frac{2 \overline{y}}{\left(1 - q^*\right) \left[\left(\theta_H - \theta_L\right) - \left(1 - \theta_L\right) \gamma\right] - q^* \left(1 - \theta_H\right) \gamma}
$$

4.4.1 The effect of a change in gain of economic success to equilibrium

This subsection investigates the the effect of change in the gain of economic success or income inequality to equilibrium. Figure 4.6 and Figure 4.7 depict images in each cases: high empathy and low empathy.

Proposition 8 *The effect of change in the gain of economic success or income inequality to equilibrium is*

$$\frac{\partial q^*}{\partial \overline{y}} = \frac{\phi}{-\left(\frac{\partial d_{1t}}{\partial q_t} - \frac{\partial d_{0t}}{\partial q_t}\right)} > 0$$

$$\frac{\partial d_{1t}^*}{\partial \overline{y}} = \phi\left(1-q^*\right) + \frac{dd_1\left(q^*, S\left(q^*\right), \overline{y}\right)/dq}{\frac{dd^*\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq} - \frac{dd^*\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq}} \geq 0,$$

$$\frac{\partial d_{0t}^*}{\partial \overline{y}} = \frac{d_0\left(q^*, S\left(q^*\right), \overline{y}\right)/dq}{\frac{dd_0\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq} - \frac{dd_1\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq}} - \phi q^* \gtreqqless 0$$

Proof. The total derivative of the condition:

$$\begin{aligned} d_1\left(q^*, S\left(q^*\right), \overline{y}\right) &= d_1\left(q^*, S\left(q^*\right), \overline{y}\right)\phi\left(1-q^*\right)\left\{\overline{y} + \left[\left(\theta_H - \theta_L\right) - \left(1-\theta_L\right)\gamma\right]S\left(q^*\right)\right\} \\ &= \phi q^*\left[\left(1-\theta_H\right)\gamma S\left(q^*\right) - \overline{y}\right] \end{aligned}$$

yields

$$\frac{\partial d_1\left(q^*, S\left(q^*\right), \overline{y}\right)}{\partial \overline{y}} d\overline{y} + \frac{d_1\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq} dq^* = \frac{\partial d_0\left(q^*, S\left(q^*\right), \overline{y}\right)}{\partial \overline{y}} d\overline{y} + \frac{d_0\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq} dq^*,$$

where

$$\frac{dd_\tau\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq} = \frac{\partial d_\tau\left(q^*, S\left(q^*\right), \overline{y}\right)}{\partial q} + \frac{\partial d_\tau\left(q^*, S\left(q^*\right), \overline{y}\right)}{\partial S}\frac{\partial S\left(q^*\right)}{\partial q},$$

Hence,

$$\begin{aligned}
\frac{dq^*}{d\overline{y}} &= \frac{\frac{\partial d_1(q^*,S(q^*),\overline{y})}{\partial \overline{y}} - \frac{\partial d_0(q^*,S(q^*),\overline{y})}{\partial \overline{y}}}{\frac{dd_0(q^*,S(q^*),\overline{y})}{dq} - \frac{dd_1(q^*,S(q^*),\overline{y})}{dq}} \\
&= \frac{\phi\left(1 - q^* + q^*\right)}{\frac{d(1-q^*),\overline{y})}{dq} - \frac{dd_1(q^*,S(q^*),\overline{y})}{dq}} \\
&= \frac{\phi\left(1 - q^*\right), \overline{y})}{dq} - \frac{dd_1\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq} > 0
\end{aligned}$$

Next,

$$\begin{aligned}
\frac{d\left[d_1\left(q^*, S\left(q^*\right), \overline{y}\right)\right]}{d\overline{y}} &= \frac{\partial d_1\left(q^*, S\left(q^*\right), \overline{y}\right)}{\partial \overline{y}} + \frac{dd_1\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq}\frac{dq^*}{d\overline{y}} \\
&= \phi\left(1 - q^*\right) + \phi\frac{dd_1\left(q^*, S\left(q^*\right), \overline{y}\right)/dq}{\frac{dd_0(q^*,S(q^*),\overline{y})}{dq} - \frac{dd_1(q^*,S(q^*),\overline{y})}{dq}} \\
&= \phi\frac{d_0\left(q^*, S\left(q^*\right), \overline{y}\right)/dq}{\frac{dd_0(q^*,S(q^*),\overline{y})}{dq} - \frac{dd^2(q^*,S(q^*),\overline{y})}{dq}} - \phi q^*
\end{aligned}$$

Therefore, if

$$\frac{dd_0\left(q^*, S\left(q^*\right), \overline{y}\right)/dq}{\frac{d_0(q^*,S(q^*),\overline{y})}{dq} - \frac{d_1(q^*,S(q^*),\overline{y})}{dq}} > q^*,$$

then, $d\left[d_1\left(q^*, S\left(q^*\right), \overline{y}\right)\right]/d\overline{y} = d\left[d_0\left(q^*, S\left(q^*\right), \overline{y}\right)\right]/d\overline{y} > 0$. This condition is more likely to hold with larger $d_1\left(q^*, S\left(q^*\right), \overline{y}\right)/dq$. When

$$\frac{dd_1\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq} > 0,$$

then, the relationship

$$\frac{dd_0\left(q^*, S\left(q^*\right), \overline{y}\right)/dq}{\frac{d_0(q^*,S(q^*),\overline{y})}{dq} - \frac{d_1(q^*,S(q^*),\overline{y})}{dq}} > q^* > 1,$$

implies that the condition always holds. ■

The effect on the equilibrium proportion of sensitive types in a change with the gain of economic success or income inequality is positive because the sensitive type's incentive to socialise is increased and that of the insensitive is increased. These changes in incentives to socialise are generated from the fact that, with greater income inequality, the greater the children's effort.

4.4.2 The effect of a change in empathy to equilibrium

Next, the effect of change in the intensity of empathy to equilibrium is as given.

Proposition 9 *The effect of a change in the intensity of empathy to equilibrium is*

$$\frac{\partial q^*}{\partial \gamma} = \frac{\frac{\partial d_{1t}}{\partial \gamma} - \frac{\partial d_{0t}}{\partial \gamma}}{-\left(\frac{\partial d_{1t}}{\partial q_t} - \frac{\partial d_{0t}}{\partial q_t}\right)}$$

$$= \frac{-\phi \left[\left(1-\theta_L\right)\left(1-q^*\right) + \left(1-\theta_H\right)\right] S\left(q^*\right)}{-\left(\frac{\partial d_{1t}}{\partial q_t} - \frac{\partial d_0}{\partial q_t}\right)} < 0$$

$$\frac{d\left[d_1\left(q^*, S\left(q^*\right), \overline{y}\right)\right]}{d\gamma} = \frac{\partial d_1\left(q^*, S\left(q^*\right), \overline{y}\right)}{\partial \gamma} + \frac{dd_1\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq}\frac{dq^*}{d\gamma}$$

$$= -\phi\left(1-q^*\right)\left(1-\theta_L\right) S\left(q^*\right)$$

$$+ \frac{dd_1\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq} \frac{\frac{\partial d_{1t}}{\partial \gamma} - \frac{\partial d_{0t}}{\partial \gamma}}{\frac{dd_0\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq} - \frac{dd_1\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq}}$$

$$\geq 0$$

$$\frac{d\left[d_0\left(q^*, S\left(q^*\right), \overline{y}\right)\right]}{d\gamma} = \frac{\partial d_0\left(q^*, S\left(q^*\right), \overline{y}\right)}{\partial \gamma} + \frac{dd_0\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq}\frac{dq^*}{d\gamma}$$

$$= \phi q^*\left(1-\theta_H\right) S\left(q^*\right)$$

$$+ \frac{dd_1\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq} \frac{\frac{\partial d_{1t}}{\partial \gamma} - \frac{\partial d_{0t}}{\partial \gamma}}{\frac{dd_0\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq} - \frac{dd_1\left(q^*, S\left(q^*\right), \overline{y}\right)}{dq}}$$

$$\geq 0$$

The effect on the proportion of the sensitive type, in a steady state of change of the intensity of empathy is negative because the sensitive type's incentive to socialise is reduced and insensitive type's is increased. This change in incentive to socialise is generated by the fact that, when more parents have empathy, the more they try to protect their children from stigma.

4.5 Conclusion

This paper examined the aspiration enhancing role of stigma from low-income of poverty. I focused on the possibility that stigma gives children incentives to escaping from poverty, and I investigated the mechanism by which a culture stigmatises earning low incomes, and the stigma is passed inter-generations. I showed that the increase in income inequality immobilises family culture evolution. However, social mobility was restrained under some conditions.

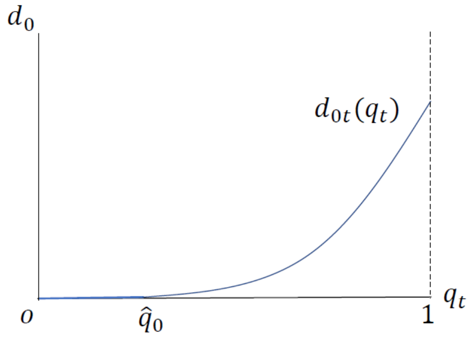

FIGURE 4.1: $d_{0t}(q_t)$

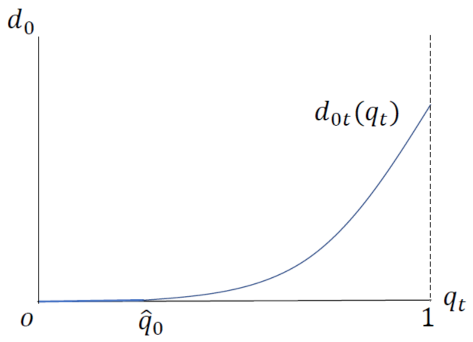

FIGURE 4.2: $d_1(q_t)$ in the case of $\gamma < \frac{\theta_H - \theta_L}{1 - \theta_L}$

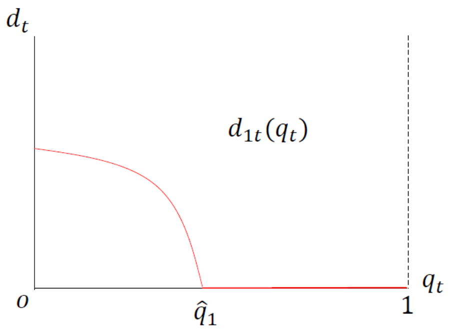

FIGURE 4.3: $d_1(q_t)$ in the case of $\gamma \geq \frac{\theta_H - \theta_L}{1 - \theta_L}$

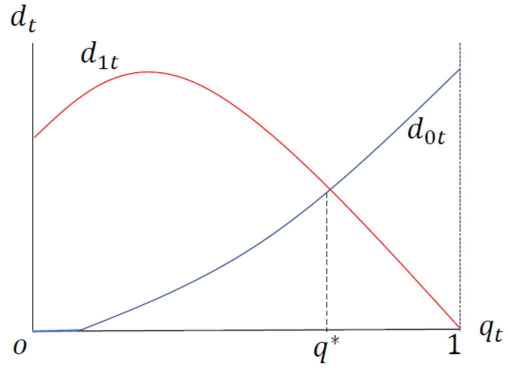

FIGURE 4.4: Steady state in the case of $\gamma < \frac{\theta_H - \theta_L}{1 - \theta_L}$

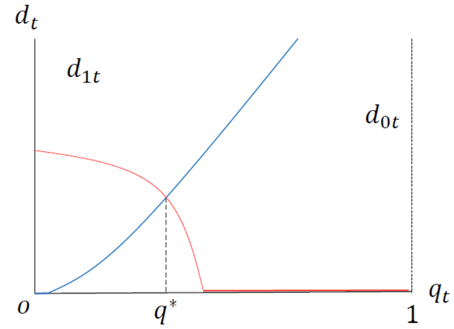

FIGURE 4.5: Steady state in the case of $\gamma \geq \frac{\theta_H - \theta_L}{1-\theta_L}$

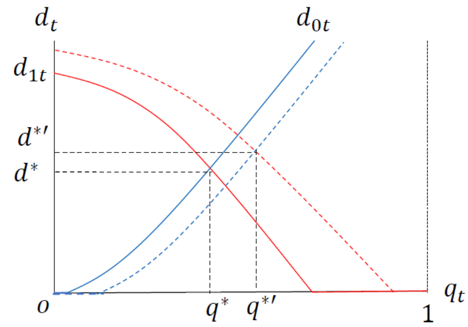

FIGURE 4.6: The effect of change in $\overline{y}$

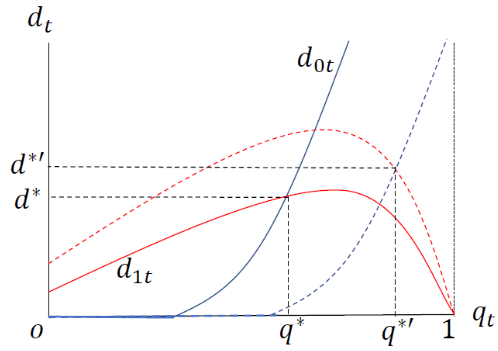

FIGURE 4.7: The effect of change in $\overline{y}$

Chapter 5

How does benefit level affect welfare stigma?: Theory and evidence from OECD panel data

Abstract

This chapter investigates the relationship between benefit level and the beneficiary ratio in minimum income guarantee program from the view of welfare stigma using theoretical analysis and empirical evidence. First, the theoretical study presents a simple stigma model to consider household decision-making regarding whether to work or take up welfare benefits. As a theoretical result, an equilibrium recipients ratio forms the inverse U-shaped curve with respect to benefit level when stigma cost is an increasing convex function with benefit level. On the other hand, an equilibrium recipients ratio appears to be the U-shaped function with benefit level when stigma cost is an increasing concave function with benefit level. Second, an empirical analysis is conducted using OECD panel data to examine which case of stigma is in keeping with the estimation result regarding the relationship between benefit level and beneficiary ratio. The empirical results are consistent with the case that stigma cost is convex function with respect to benefit level.

5.1 Introduction

To analyse welfare programs, most public economics researchers exploit a labour supply model that is based on the maximisation problem of leisure and consumption goods. This model can explain welfare fraud but not non-take-up of welfare; however, refusal to take up welfare occurs in most developed countries (Currie, 2006; Immervoll, 2009; Plueger, 2009). 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 behaviour by society or a social group. In particular, stigma is an important concept in social psychology (Major, Dovidio, and Link, 2017).

Moffitt (1983) conducted one of the earlier studies to focus on welfare stigma in economics by analysing household decision-making regarding whether to take up welfare benefits or supply labour 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 Coate (1992) presented a model where the level of welfare stigma is determined endogenously. Two models were analysed, the statistical discrimination view stigma model and the tax-resentment view stigma model. Lindbeck, Nyberg, and Weibull (1999) used a simple form of stigma cost, namely, the decreasing function of the proportion of beneficiaries. They included a political process for choosing a level of transfer and tax rate and stated the possibility that multiple equilibria could occur. Hupkau and Maniquet (2018) analysed the problem of non-take-up of welfare from the perspective of identity economics (Akerlof and Kranton, 2000; Rachel, 2016).

In empirical studies of welfare stigma, researchers have tried to estimate the take-up rate of welfare benefit programmes or social benefit programmes. For example, Bargain, Immervoll, and Viitamäki (2012) pointed out that non-take-up of welfare has been demonstrated in Finland. Bhargava and Manoli (2015) investigated the causes of non-take-up of welfare, conducted field experiments with the US Internal Revenue Service (IRS) and estimated fractions of welfare take-up using difference in difference estimation. They suggested that welfare stigma has a statistically significant impact on welfare take-up rates. Friedrichsen, König, and Schmacker (2018) indicated that the existence of stigma reduces the take-up rate by about 30 percents.

It is difficult to measure welfare stigma directly without micro-data or experiment. This chapter will attempt to investigate the relationship between welfare stigma and level of welfare by identifying the causal effect of benefit level on recipient ratio. We will present a simple model of welfare stigma to compare theoretical results between several functional forms of stigma cost. Moreover, we test empirically our theoretical results by using OECD panel data. Our empirical results are in line with the case that stigma cost function is convex with respect to benefit level.

This chapter is organised as follows. The next section describes the theoretical model. Section 5.3 presents to the approach used to analyse the relationship of interest, gives details of the data, and presents the empirical results. Section 5.4 concludes the paper with a discussion of our findings and suggests future research questions.

5.2 Theoretical model

This section introduces the theoretical model used to explain the relation between recipients ratio and benefit level. Each household chooses a course of action: work or take up welfare benefits. Following to Besley and Coate (1992), household utility is as follows:

$$\begin{cases} u(w) - \theta & \text{if working} \\ u(b) - s & \text{if taking-up,} \end{cases}$$

where w is the wage rate distributed as a continuous function $F(w)$, i.e., $w \sim F(w)$, b is a level of welfare benefit, θ is disutility of labour, s is a level of stigma cost. $u(\cdot)$ is utility from consumption, where $u'(\cdot) > 0$, and $u''(\cdot) \geq 0$. Thus, when an individual chooses to work, they can enjoy consumption w but suffers a disutility from labour. On the other hand, when a person chooses to take up welfare benefit, they can enjoy a consumption without a disutility from labour but incur a psychological cost driven by welfare stigma. Each household makes a decision under such a trade-off. Each person chooses their individual action, taking a level of stigma cost, s, as given. We define the critical level of wage rate $\hat{w}$ as follows:

$$u(\hat{w}) - \theta = u(b) - s.$$

Then, households for whom w is less than or equal to $\hat{w}$ will choose to take-up welfare benefit. In contrast, households whose wage rate is larger than $\hat{w}$ will choose to work. Therefore, a recipients ratio, r, can be determined:

$$r = F(\hat{w}).$$

we introduce the formation of stigma cost used in this model. This research assumes that stigma cost as follows:

$$s = s_0 + s_1(b), \;\; s_1'(\cdot) > 0.$$

where s_0 is fixed component of stigma from participation in the program, $s_1(\cdot)$ is variable component of stigma with respect to benefit level, $s_1'(\cdot)$. This assumption implies that the stigma cost rises with the level of welfare benefits recipients received. It is inspired by Besley and Coate (1992)'s taxpayer resentment view model[1].

This assumption implies that the stigma cost rises with the level of welfare benefits received, and the rate of this increase also rises with level of welfare benefits.

For simplicity, we specify the utility function of consumption as the cumulative distribution function of wage and stigma cost function as follows:

$$u(c) = \alpha c, \;\; \alpha > 0,$$
$$s_1(b) = b^{\sigma}, \;\; \sigma > 0,$$
$$w \sim \mathcal{U}[0, \overline{w}].$$

That is, stigma cost function is convex when $\sigma > 1$ and concave when $\sigma \in (0,1)$. The above settings and assumptions yield the following results with regard to the critical wage $\hat{w}$ and recipients ratio r

$$\begin{aligned} \hat{w} &= u^{-1}[u(b) - s + \theta] \\ &= \frac{1}{\alpha \overline{w}}[\alpha b - s_0 - b^{\sigma} + \theta], \\ r &= F(\hat{w}) = \frac{\hat{w}}{w}. \end{aligned}$$

Therefore, an equilibrium recipient ratio can be obtained as follows:

$$r^* = \frac{1}{\alpha w}[\alpha b + \theta - s_0 - b^{\sigma}].$$

[1] The Besley and Coate (1992)'s model analysed a level of stigma cost as endogenous variable while we assume that stigma cost is just function of benefit level for the analytical simplicity.

We are interested in the relationship between the recipient ratio and level of benefit. The following proposition shows the effect of a change in benefit level on the equilibrium recipient ratio:

Proposition 10 *When $\sigma > 1$,*

$$\frac{\partial r^*}{\partial b} = \frac{1}{\alpha}(\alpha - \sigma b^{\sigma-1}) \geq 0 \quad \text{if } b \leq \left(\frac{\alpha}{\sigma}\right)^{\frac{1}{\sigma-1}},$$
$$\frac{\partial r^*}{\partial b} = \frac{1}{\alpha}(\alpha - \sigma b^{\sigma-1}) < 0 \quad \text{if } b > \left(\frac{\alpha}{\sigma}\right)^{\frac{1}{\sigma-1}}.$$

When $\sigma \in (0, 1)$,

$$\frac{\partial r^*}{\partial b} = \frac{1}{\alpha}(\alpha - \sigma b^{\sigma-1}) \leq 0 \quad \text{if } b \leq \left(\frac{\alpha}{\sigma}\right)^{\frac{1}{\sigma-1}},$$
$$\frac{\partial r^*}{\partial b} = \frac{1}{\alpha}(\alpha - \sigma b^{\sigma-1}) > 0 \quad \text{if } b > \left(\frac{\alpha}{\sigma}\right)^{\frac{1}{\sigma-1}}.$$

When $\sigma = 1$,

$$\frac{\partial r^*}{\partial b} = \frac{1}{\alpha}(\alpha - 1) < 0 \quad \text{if } \alpha < 1,$$
$$\frac{\partial r^*}{\partial b} = \frac{1}{\alpha}(\alpha - 1) = 0 \quad \text{if } \alpha = 1,$$
$$\frac{\partial r^*}{\partial b} = \frac{1}{\alpha}(\alpha - 1) > 0 \quad \text{if } \alpha > 1.$$

The relationship between r^* and b is divided into five cases. If $\sigma > 1$, an equilibrium recipients ratio r^* is upper convex function of benefit level b, as shown in Figure 5.1. To explain why $r^*(b)$ is an inverse U-shaped curve, consider the case of sufficiently low benefit level, i.e., $b \leq (\frac{\alpha}{\sigma})^{\frac{1}{\sigma-1}}$. In this case, the marginal utility of a rising benefit level is higher than the marginal stigma cost. Thus, the equilibrium recipient ratio increases. On the other hand, when the benefit level is sufficiently high, $b > (\frac{\alpha}{\sigma})^{\frac{1}{\sigma-1}}$, then the marginal utility is lower than the marginal stigma cost, and the equilibrium recipients ratio therefore decreases. Figure 5.2 shows the case of $\sigma \in (0, 1)$. In this case, $r^*(b)$ is U-shaped function with benefit level. Figures 5.3, 5.4 and 5.5 show cases of $\sigma = 1$. If $\alpha > 1$, r^* is an increasing linear function with benefit level b as shown in Figure 5.3. If α equals to one, r^* is constant to level of welfare b as depicted in Figure 5.4. If $\alpha < 1$, $r^*(b)$ is an decreasing linear function with b as shown in Figure 5.5.

In this section, we analysed a simple stigma model. The analysis results showed that the equilibrium recipient ratio can be depicted as the upper convex curve with respect to a benefit level when stigma cost function is convex. In the next section, we determine which case is consistent with the real situation by using OECD panel data.

5.3 Empirical analysis

This section presents the empirical analysis conducted to clarify the relationship between the recipients ratio and the minimum income benefit level in accordance with the theoretical model in 5.2.

5.3.1 Econometric model

The panel data were analysed to investigate the causal effect of the minimum guaranteed income level on social benefit recipients. The decision to apply the panel data to our theoretical model reflects three motivations. First, a panel data model can have better prediction accuracy than 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 of society in the empirical analysis (Greene, 2012). This chapter analyses the relationship between the minimum income benefit level and social benefit recipient ratio on the basis of the baseline model:

$$y_{it} = \mathbf{x}'_{it}\boldsymbol{\beta} + \varepsilon_{it}, \tag{5.1}$$

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 ε_{it} is the disturbance term, which is distributed as $\varepsilon_{it} \sim \mathcal{N}(0, \sigma_\varepsilon^2)$. Furthermore, in equation (5.1), $i = 1, \ldots, n$ indicates the index for a country, whereas $t = 1, \ldots, T$ represents the index for time. The OLS estimation of equation (5.1) after pooling the available data is called the pooling estimation.

When we consider the country-specific heterogeneity in the disturbance term of equation (5.1), ε_{it} can be decomposed as follows:

$$\begin{aligned} y_{it} &= \mathbf{x}'_{it}\boldsymbol{\beta} + \varepsilon_{it} \\ \varepsilon_{it} &= \alpha_i + \eta_{it}, \end{aligned} \tag{5.2}$$

where α_i is the error depending on the country i and $\eta_{it} \sim$ i.i.d. $\mathcal{N}(0, \sigma_\eta^2)$ is the stochastic disturbance term. Equation (5.2) can be considered a one-way error component model (Baltagi, 1984) because it decomposes the disturbance term ε_{it} into the error based on the individual heterogeneity and the stochastic error. The model in equation (5.2) 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 (5.1), the disturbance term can be further decomposed to incorporate heterogeneity in time:

$$\begin{aligned} y_{it} &= \mathbf{x}'_{it}\boldsymbol{\beta} + \varepsilon_{it} \\ \varepsilon_{it} &= \alpha_i + \lambda_t + \eta_{it}, \end{aligned} \tag{5.3}$$

where λ_t is the error depending on the time t. Equation (5.3), 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 (5.2),the model of equation (5.3) 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 chapter 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).

5.3.2 Data

This section proposes the detail of our dataset used for estimation of the panel data models introduced in Section 5.3.1. 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 number of social benefit recipients were retrieved from 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 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 unknown 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 Labour Force Statistics.

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

5.3.3 Result

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

Table 5.1 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 5.1 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 5.2 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 5.3 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.

Table 5.4 shows the estimation results based on the data introduced in Section 5.3.2. 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 5.4 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.

This upper convex relationship has the following implications. When the benefit level is sufficiently low, the marginal utility of an increase in the benefit level is higher than the marginal stigma cost. On the other hand, when the benefit level is sufficiently high, the marginal utility is lower than the marginal stigma cost.

The empirical results presented in this section have demonstrated the existence of an upper convex relation between the benefit level and the recipient ratio. These empirical results are consistent with the case that stigma cost function is convex with benefit level as described in Case 1 of Proposition 10.

5.4 Conclusion

This research has tried to determine whether a country benefit level affects welfare stigma by using a combined theoretical and empirical approach. The theoretical prediction can be broken down into five cases. The empirical analysis tested the theory predictions using OECD macro-panel data. Our obtained empirical results are consistent with the theoretical model case of convex stigma with respect to benefit level.

This chapter focused on the threshold model in theoretical analysis because we conducted the empirical analysis using macro data. However, the threshold model was only able to analyse the extensive margin, not the intensive margin. Accordingly, future research should analyse the stigma model, including the intensive margin (Saez, 2002).

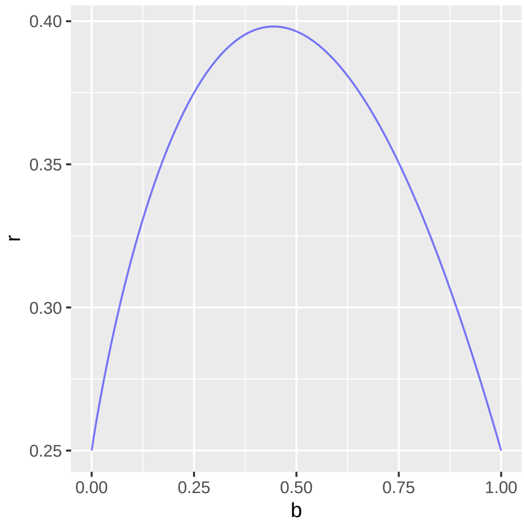

FIGURE 5.1: Case 1, $\sigma > 1$

Notes: $\alpha = 1$, $\theta = 0.3$, $\sigma = 1.5$.

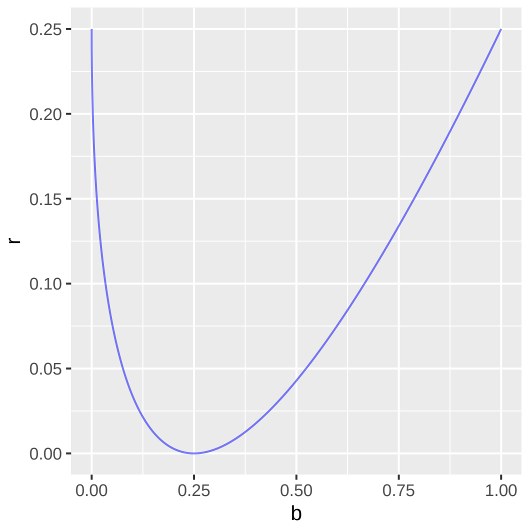

FIGURE 5.2: Case 2, $\sigma < 1$

Notes: $\alpha = 1$, $\theta = 0.3$, $\sigma = 0.5$.

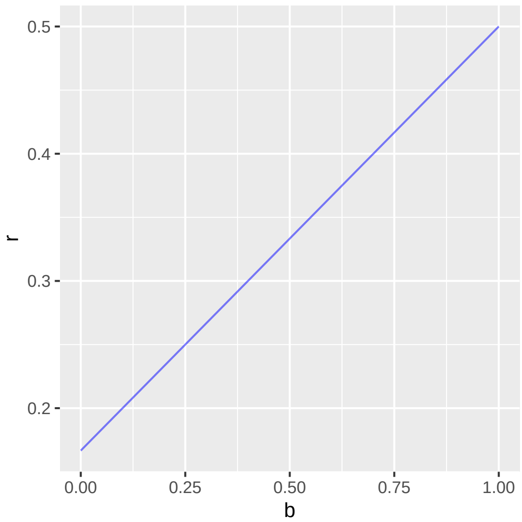

FIGURE 5.3: Case 3, $\sigma = 1$ and $\alpha > 1$

Notes: $\alpha = 1.5$, $\theta = 0.3$, and $\sigma = 1$.

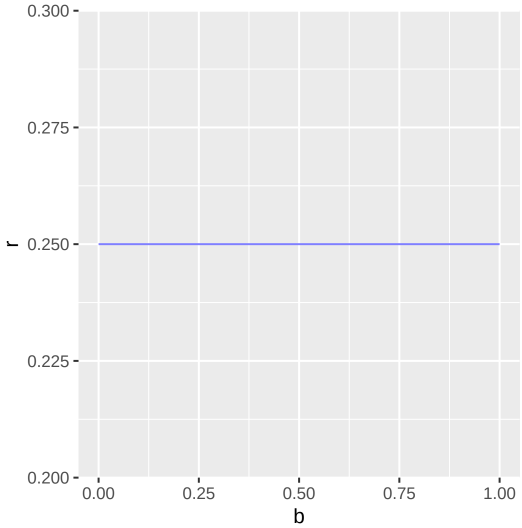

FIGURE 5.4: Case 4, $\sigma = 1$ and $\alpha = 1$

Notes: $\alpha = 1$, $\theta = 0.3$, and $\sigma = 1$.

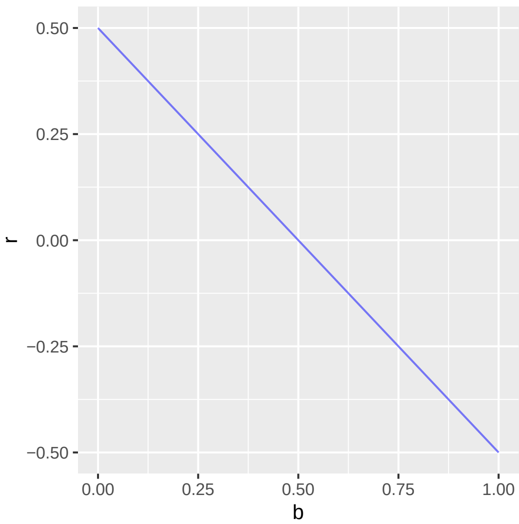

FIGURE 5.5: Case 5, $\sigma = 1$ and $\alpha < 1$

Notes: $\alpha = 0.5$, $\theta = 0.3$, and $\sigma = 1$.

TABLE 5.1: 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 5.2: 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 5.3: 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 5.4: 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.

Chapter 6

Conclusion

This dissertation provides theoretical and empirical analysis on welfare stigma. Some key insights and results contribute to the understanding of relationship between stigma and household's decision-making.

The second chapter presented the extended version of statistical discrimination stigma model. The model had been developed by Besley and Coate (1992). There existed two types in the model: the needy type and the non-needy type. The needy type people could not work even if they hope to do. On the other hand, the non-needy type people could work even when they would to do. Stigma cost was determined by the ratio of the non-needy type to the needy type in recipients. In particular, they had assumed that stigma cost is a decreasing function with the ratio. Their results had indicated the occurrence of welfare fraud. However, the needy type had been assumed to take-up welfare regardless of level of stigma cost. In general, people like needy type are thought to be influenced by stigma in their decision-makings. Moreover, not taking-up welfare by needy type corresponds to ro-kyu which means not taking-up welfare by eligible poor people. To solve these problems and limitations in previous research, I extended the statistical stigma model to endogenise the decision-making process for the needy type. As a result, multiple equilibrium could occur: one equilibrium was 'high stigma, serious ro-kyu and restrained welfare fraud' and the other was 'low stigma, restrained ro-kyu, serious welfare fraud'. Multiple equilibria were likely to occur when the needy type's elasticity of stigma sensitiveness to the ratio of the needy type to the non-needy type in the pool of recipients in welfare benefit. The comparative static analysis indicated that an increase in the benefit level was likely to reduce take-up of welfare benefits in needy types. I demonstrated this result through a simple equation, linking two types of elasticity for needy and non-needy types: the needy type's elasticity of stigma sensitiveness to the ratio between types in recipients and the non-needy type's elasticity of material utility to benefit level were sufficiently large.

The third chapter studied the stigma model of relative income. The model was similar to the taxpayer resentment model (Besley and Coate, 1992) in the point that non-recipients (workers, taxpayers or capitalists). The model differed from Besley and Coate (1992)'s taxpayer resentment model in that the stigmatisee can become the stigmatiser. Moreover, we used relative income mechanism to formulate stigma which was increasing function with average of the ratio of benefit level to each working wage. As a result, multiple equilibria could occur, high stigma and low stigma, contrary to the proportion of Besley and Coate (1992)'s model. It was because there exists feedback effect. In the comparative static analysis, the result indicated that the possibility that the number of recipients declined in case of a negative macroeconomic shock since there existed the negative indirect effect on recipients, which decreased in case of increase in stigma from increased resentment due to a shift of working income distribution.

The fourth chapter examined the aspiration enhancing role of stigma from low-income of poverty. This chapter focused on the possibility that stigma gives children incentives to escaping from poverty. I investigated the mechanism by which a culture stigmatises earning low incomes and the stigma is passed inter-generation. I showed that the increase in income inequality immobilises family culture evolution. However, social mobility was restrained under some conditions.

The fifth chapter tried to determine whether a country benefit level affects welfare stigma by using a combined theoretical and empirical approach. The theoretical prediction were broken down into five cases. The empirical analysis tested the theory predictions using OECD macro-panel data. Our obtained empirical results were consistent with the theoretical model case of convex stigma with respect to benefit level. This chapter focused on the threshold model in theoretical analysis because we conducted the empirical analysis using macro data. However, the threshold model was only able to analyse the extensive margin, not the intensive margin. Accordingly, future research should analyse the stigma model, including the intensive margin (Saez, 2002).

Bibliography

[1] Akerlof, G. A. and Kranton, R. E. "Economics and identity". *Quarterly Journal of Economics* 115.3 (2000), pp. 715–753.

[2] Arellano, M. "Computing Robust Standard Errors for Within-groups Estimators". *Oxford Bulletin of Economics and Statistics* 49.4 (1987), pp. 431–434.

[3] Atkinson, A. B. *Public Economics in Action: The basic income/Flat Tac Proposal*. Oxford: Oxford Uniersity Press, 1995.

[4] Baltagi, B. H. "A Monte Carlo Study for Pooling Time Series of Cross-Section Data in the Simultaneous Equations Model". *International Economic Review* 25.3 (1984), pp. 603–624.

[5] Baltagi, B. H. *Econometric analysis of panel data*. John Wiley & Sons, 2008.

[6] Bargain, O., Immervoll, H., and Viitamäki, H. "No claim, no pain. Measuring the non-take-up of social assistance using register data". *Journal of Economic Inequality* 10.3 (2012), pp. 375–395.

[7] Becker, G. S. "A Theory of the Allocation of Time". *The economic journal* (1965), pp. 493–517.

[8] Benabou, R. and Tirole, J. "Belief in a just world and redistributive politics". *The Quarterly journal of economics* 121.2 (2006), pp. 699–746.

[9] Besley, T. and Coate, S. "Understanding welfare stigma: Taxpayer resentment and statistical discrimination". *Journal of Public Economics* 48.2 (1992), pp. 165–183.

[10] Besley, T. "Aspirations and the political economy of inequality". *Oxford Economic Papers* 69.1 (2016), pp. 1–35.

[11] Bhargava, S. and Manoli, D. "Psychological Frictions and the Incomplete Take-Up of Social Benefits: Evidence from an IRS Field Experiment". *American Economic Review* 105.11 (2015), pp. 3489–3529.

[12] Bisin, A. and Verdier, T. "The economics of cultural transmission and the dynamics of preferences". *Journal of Economic Theory* 97.2 (2001), pp. 298–319.

[13] Bisin, A. and Verdier, T. ""Beyond the melting pot": cultural transmission, marriage, and the evolution of ethnic and religious traits". *The Quarterly Journal of Economics* 115.3 (2000), pp. 955–988.

[14] Blank, R. M. and Ruggles, P. "When do women use aid to families with dependent children and food stamps?" *Journal of Human Resources* 31.1 (1996), pp. 57–89.

[15] Blumkin, T., Margalioth, Y., and Sadka, E. "Welfare Stigma Re-Examined". *Journal of Public Economic Theory* 17.6 (2015), pp. 874–886.

[16] Currie, J. "The Take-up of Social Benefits". *Public Policy and the Income Distribution*. Ed. by Auerbach A. J., C. D. and Quigley, J. M. New York: Russell Sage Foundation, 2006, pp. 80–148.

[17] Duclos, J. "Modelling the take-up of state support". *Journal of Public Economics* 58.3 (1995), pp. 391–415.

[18] Fitzpatrick, T. *Freedom and security: An introduction to the basic income debate*. Hampshire: Macmillan Press, 1999.

[19] Friedman, M. *Capitalism and freedom*. Chicago: University of Chicago Press, 1962.

[20] Friedrichsen, J., König, T., and Schmacker, R. "Social image concerns and welfare take-up". *Journal of Public Economics* 168 (2018), pp. 174–192.

[21] Ghatak, M. and Maniquet, F. "Universal Basic Income: Some Theoretical Aspects". 2019.

[22] Goffman, E. *Stigma: Notes on the Management of Spoiled Identity*. Prentice-Hall, 1963.

[23] Greene, W. H. *Econometric analysis*. 7th. Pearson, 2012.

[24] Hausman, J. A. "Specification Tests in Econometrics". *Econometrica* 46.6 (1978), pp. 1251–1271.

[25] Heckman, J. J. "Dummy endogenous variables in a simultaneous equation system". *Econometrica* 46 (1977), pp. 931–959.

[26] Honda, Y. “Testing the Error Components Model with Non-Normal Disturbances”. *The Review of Economic Studies* 52.4 (1985), pp. 681–690.

[27] Horan, P. M. and Austin, P. L. “The social bases of welfare stigma”. *Social Problems* 21 (1974), pp. 648–657.

[28] Hupkau, C. and Maniquet, F. “Identity, non-take-up and welfare conditionality”. *Journal of Economic Behavior and Organization* 147 (2018), pp. 13–27.

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

[30] Kayser, H. and Frick, J. R. “Take it or leave it:(non-) take-up behavior of social assistance in Germany”. *DIW Discussion Papers* 210 (2000).

[31] Lee, L. F. “Identification and estimation in binary choice models with limited (censored) dependent variables”. *Econometrica* 47 (1979), pp. 977–996.

[32] Lewis, O. *The Children of Sanchez: Autobiography of a Mexican Family.* Random House, 1966.

[33] Lindbeck, A., Nyberg, S., and Weibull, J. W. “Social Norms and Economic Incentives in the Welfare State”. *Quarterly Journal of Economics* 114.1 (1999), pp. 1–35.

[34] Major, B., Dovidio, J. F., and Link, B. G. *The oxford handbook of stigma, discrimination, and health.* Oxford: Oxford Univ Press, 2017.

[35] Meltzer, A. H. and Richard, S. F. “A Rational Theory of the Size of Government”. *Journal of Political Economy* 89.5 (1981), pp. 914–927.

[36] Moffitt, R. “An economic model of welfare stigma”. *American economic review* 73.5 (1983), pp. 1023–1035.

[37] Moffitt, R. “The negative income tax and the evolution of US welfare policy”. *Journal of Economic Perspectives* 17.3 (2003), pp. 119–140.

[38] OECD. “OECD.Stat”. *(Accessed on 1 February 2019)* (2019).

[39] Piketty, T. “Social mobility and redistributive politics”. *The Quarterly journal of economics* 110.3 (1995), pp. 551–584.

[40] Plueger, D. “Earned Income Tax Credit participation rate for tax year 2005”. *IRS Research Bulletin* (2009).

[41] Rachel, K. “Identity Economics 2016: Where Do Social Distinctions and Norms Come From?” *American Economic Review: Papers & Proceedings* 286.6 (2016), pp. 139–146.

[42] Rainwater, L. “Stigma in income-tested programs”. *Income-Tested Transfer Programs* (1982), pp. 19–65.

[43] Riphahn, R. T. "Rational poverty or poor rationality? The take-up of social assistance benefits". *Review of Income and Wealth* 47.3 (2001), pp. 379–398.

[44] Robins, P. K. "A comparison of the labor supply findings from the four negative income tax experiments". *Journal of Human Resources* 20 (1985), pp. 567–582.

[45] Saez, E. "Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses". *The Quarterly Journal of Economics* 117.3 (2002), pp. 1039–1073.

[46] Shayo, M. "A model of social identity with an application to political economy: Nation, class, and redistribution". *Ameican Political Science Review* 103.2 (2009), pp. 147–174.

[47] Smeeding, T., Erikson, R., and Jäntti, M. *Persistence, privilege, and parenting: The comparative study of intergenerational mobility*. Russell Sage Foundation, 2011.

[48] Tachibanaki, T and Urakawa, K. *A study on poverty in Japan*. Tokyo: The University of Tokyo Press, 2006.