Panel Data Research Center at Keio University
DISCUSSION PAPER SERIES

DP2016-006
February, 2017

Spatial Probit Analysis on Welfare Stigma: Evidence from Japan

Masato TAKAHASHI*

【Abstract】
This paper empirically examines welfare stigma in Japan using the spatial autoregressive and error probit models based on the Japan Household Panel Survey (JHPS/KHPS). The empirical analysis finds evidence that favours the existence of internal type of traditional stigma; however, no evidence suggests the existence of external type of traditional stigma. This paper contributes to the existing literature in two ways: first, the public assistance system in Japan is for the first time empirically examined from the perspective of welfare stigma; second, it shows that internal and external types of traditional stigma can be simultaneously examined using the spatial probit models.

* Faculty of Economics, University of Tokyo Candidate for the B.A. in Economics

Panel Data Research Center at Keio University
Keio University

Spatial Probit Analysis on Welfare Stigma: Evidence from Japan*

Masato TAKAHASHI†

Faculty of Economics
University of Tokyo

January 2017

Abstract

This paper empirically examines welfare stigma in Japan using the spatial autoregressive and error probit models based on the Japan Household Panel Survey (JHPS/KHPS). The empirical analysis finds evidence that favours the existence of internal type of traditional stigma; however, no evidence suggests the existence of external type of traditional stigma. This paper contributes to the existing literature in two ways: first, it is the first study to empirically examine welfare stigma in the public assistance system in Japan; second, it shows that internal and external types of traditional stigma can be examined at the same time using the spatial probit models.

Keywords: Public assistance, Spatial probit model, Welfare stigma
JEL Classification: C21, I38

*I would like to thank Professor Masayoshi Hayashi for his guidance and advice. I am also grateful to Dr Kei Kawakami for his encouragement to work on this independent research project. The JHPS/KHPS data was provided courtesy of the Panel Data Research Center at Keio University. All errors are mine.

†Candidate for the B.A. in Economics

1 Introduction

The Japanese social security system's raison d'être is to ensure the minimum level of living not only by providing people in need with social insurance and public assistance but also by promoting social welfare and public health. This paper aims to empirically examine the existence of welfare stigma in the public assistance system in Japan. For the purpose of better understanding the role of today's public assistance system, this introduction first provides a brief overview of the social security system.

In the field of industrial relations, it is often argued that the social security system is a by-product of the long-term employment system that emerged in the 20$^{\text{th}}$ century. In fact, according to the Ministry of Health, Labour and Welfare (2000), the English expression 'social security' first appeared in the Social Security Act enacted in 1935 in the United States. In general, the social security system consists of the social insurance system, the public assistance system, and various welfare-related services[1]. However, these components do not share equal importance; it is the social insurance system that has been the core of the social security system throughout the 20$^{\text{th}}$ century.

The essential distinction between social insurance and public assistance is that the former is an instrument for poverty prevention while the latter is that for poverty relief. In this respect, the superiority of social insurance within the social security system is not a surprising consequence considering that poverty prevention was widely adopted as the basic principle for poverty issues at the time when the social security system was formed.

As social insurance became the core of the social security system, public assistance was given its role as the last resort, or, in this particular context, the last safety net. At this point, however, it has yet to be explained why poverty relief was not chosen as a primary instrument for reducing poverty in the first place. The key to understanding this seemingly arbitrary consequence is the existence of welfare stigma that typically accompanies public assistance.

[1]It should be noted that even though such systems and services did not become interdependent until the 20$^{\text{th}}$ century, respective components of the social security system had already existed prior to its establishment; for example, the social insurance system was already available in the 1880s in Germany (Martin and Weaver, 2005).

The existence of welfare stigma means that there are people who choose not to receive public assistance despite satisfying the eligibility criteria. Many among those eligible make such a decision because they fear that they would be negatively labelled, disapproved of, or shamed in public once they participate in a public assistance programme. Given such a negative perception of public assistance, it is not surprising that this lifeline has never been viewed favourably, not even when the social security system was established.

Even though the public assistance system is designed to complement the social insurance system, the low take-up rate of public assistance due to welfare stigma is problematic in that it creates inefficiency as a results of preventing the social security system from functioning as expected. This implies that the social security system is incomplete unless the public assistance system works properly. Hence, not only has welfare stigma been of academic interest in sociology and economics for the past several decades but also its reduction has been considered to be one of the key policy issues.

Various economic studies have found evidence of welfare stigma; for example, Moffitt (1983) and Kayser and Frick (2000) indicate that welfare stigma exists in the United States and Germany, respectively. Meanwhile, to the author's knowledge, there has been no empirical research on welfare stigma in Japan.

Yet the lack of studies on welfare stigma in Japan does not imply that public assistance is not controversial or that the issue of welfare stigma is non-existent in Japan. On the contrary, the poverty rate in Japan has been rising and the demand for public assistance has also been increasing. Moreover, notwithstanding the rising public assistance rate in Japan, Fujisawa (2008) suggests that the take-up rate remains very low.

In light of this fact, this paper empirically examines the existence of welfare stigma in Japan using the spatial probit models based on the Japan Household Panel Survey (JHPS/KHPS). More specifically, the existence of both internal and external types of traditional stigma is examined at the same time by estimating the spatial autoregressive and error probit models using the Bayesian MCMC approach. The JHPS/KHPS is a nationwide household panel survey that has been conducted by the Panel Data Research Center at Keio University since 2009.

The spatial autoregressive probit estimation supports the existence of internal type of traditional stigma; however, it finds no evidence that suggests the existence of external type of traditional stigma. The spatial error probit estimation does not confirm the presence of spatially correlated shocks, either. Therefore, as far as the empirical analysis of this paper is concerned, no spatial effect exists in the public assistance system in Japan.

This paper contributes to the existing literature in two ways: first, it marks the first attempt to empirically examine the existence of welfare stigma in Japan; second, by conducting a spatial probit analysis on welfare stigma, it fills the gap in the existence literature that lacks an empirical framework that incorporates both internal and external types of traditional stigma.

The organisation of the rest of the paper is as follows. Section 2 reviews the relevant literature in the field of sociology and economics. Section 3 outlines the historical and systematic background of the public assistance system in Japan. Section 4 presents the model. Section 5 explains the data and the methodology. Section 6 shows the results. Section 7 concludes. All tables and figures are presented at the end of the paper.

2 Literature Review

2.1 Studies in Sociology

Sociological studies on welfare stigma are significant for two reasons: first, they contribute to the development of a theoretical foundation; and second, they open up the possibility of economic studies on the subject. This sub-section outlines a sociological background on the phenomenon of welfare stigma based on Goffman (1963), Spicker (2011), and Locke (1791).

To begin with, the phenomenon of social stigma can be associated with various ideas, e.g. physical disabilities, mental illnesses, sexual orientation. Hence, it is difficult to discuss welfare stigma, which only represents one aspect of social stigma, without referring to the study of social stigma per se. Goffman (1963) is regarded as one of the most profound studies on social stigma and makes a broad definition using various expressions such as ‘a

failing, a shortcoming, a handicap'. Many later studies in sociology as well as in economics rely on Goffman (1963) for the definition of social stigma; for instance, Besley and Coate (1992) argues that the statistical discrimination against those who receive social welfare occurs as a result of people around them perceiving 'blemishes of individual character' (Goffman, 1963) on the recipients, whether intentionally or unintentionally.

Given this basic understanding of social stigma, Spicker (2011), which was originally published in 1984, particularly focuses on the historical development of welfare stigma. Spicker (2011) argues that welfare stigma originated after the enactment of the Act for the Relief of the Poor 1601, which is commonly known as the Elizabethan Poor Law. While the 1601 Poor Law is often attributed historical significance for the establishment of a first instrument for poverty relief in history, it is less well known that it led to the creation of welfare stigma. This being the case, what aspect of the Old Poor Law[2] was responsible for such outcome?

Locke (1791), which offers an insight into this question, is the first literature to describe a phenomenon that will later be called welfare stigma, according to Spicker (2011)[3]. The significance of Locke (1791) is twofold. First, it reveals that the *dualism* of welfare stigma was already in existence in the 17$^{\text{th}}$ century[4]. Welfare stigma has two distinct effects: first, it discourages those eligible for social welfare from actually claiming such benefits; and second, it discourages those ineligible from falsely applying for various social welfare programmes. In a recent study, Blumkin et al. (2015) refers to welfare stigma of the former kind as 'traditional stigma' and that of the latter kind as 'statistical stigma'. Considering that Blumkin et al. (2015) is the first study to use such terminology, it is even more remarkable that the idea of welfare stigma being dichotomous can be traced back to the 17$^{\text{th}}$ century.

Second, it implies that statistical stigma was more pressing than traditional stigma in

[2]The Old Poor Law is a general term referring to poverty-relief-related legislation enacted between 1601 and 1834, such as the 1601 Poor Law, the Settlement Act (1662), the Workhouse Test Act (1723).

[3]Since sociology was not established – at least in the modern sense – when it was published, it might be questionable whether Locke (1791) should be covered in this sub-section; nevertheless, it is referred to here due to its high contextual relevance to the topic of welfare stigma.

[4]Although the book was published in 1791 with some notes added by the Editor, the original manuscript was written in 1697.

the 17$^{\text{th}}$ century. This implication can be inferred from the following account by the Editor that the recipients were obliged to wear a badge to prevent paupers from unnecessarily applying for relief. As wearing a badge began to be perceived as a stigma, however, the poverty relief system under the Old Poor Law became less and less effective. What is significant about this implication is that it suggests that the situation in the 20$^{\text{th}}$ century, where traditional stigma is considered to be more pressing than statistical stigma, is quite contrary to that in the 17$^{\text{th}}$ century.

Spicker (2011) claims that even after the Old Poor Law was abolished, the idea of welfare stigma remained. Then, how did the reversal of priority between statistical stigma and traditional stigma occur? This question could be answered by thinking that it was only traditional stigma that remained after the end of the Old Poor Law while statistical stigma gradually disappeared. Although this somewhat mysterious occurrence seems worthy of further consideration, it will not be further discussed in this sub-section as it deviates from the aim of this paper.

In summary, Goffman (1963) and Spicker (2011) are regarded to as having made important contributions in developing the theoretical foundation of welfare stigma from a sociological perspective; consequently, they have influenced not only subsequent sociological studies but also economic studies of welfare stigma.

2.2 Studies in Economics

Economic studies on welfare stigma, which started in the 1970s, have not only provided the empirical support for sociological studies but have also made significance contributions to a better understanding of the mechanism of welfare stigma through quantitative analysis.

Before moving on to explaining some relevant articles, this sub-section begins by introducing the four categories of welfare stigma. See Table 1 in which the vertical axis shows whether stigma is internal or external type (to be defined shortly) and the horizontal axis shows whether it is traditional or statistical. Table 2 shows that representative economic literature on welfare stigma can be classified based on this classification in Table 1. For example, Moffitt (1983) falls into Category I as it deals with traditional stigma caused by

internal factors; similarly, Levinson and Rahardja (2004) and Manchester and Mumford (2012) belong to Category I; Besley and Coate (1992) and Lindbeck et al. (1999), Category II; and Blumkin et al. (2015), Categories I, III, and IV.

Although each category has been the focus of at least one study in the literature, as shown in Table 2, it is only Blumkin et al. (2015) that examines two or more categories. Blumkin et al. (2015) incorporates both internal and external types of statistical stigma as well as internal type of traditional stigma within its model[5]. In other words, no study has so far carried out a cross-cutting examination of both internal and external types of traditional stigma. In this respect, this paper marks the first attempt to explicitly examine internal and external types of traditional stigma at the same time.

This sub-section puts emphasis on reviewing Weisbrod (1970), Moffitt (1983), Besley and Coate (1992), and Blumkin et al. (2015); however, such relevant studies as Levinson and Rahardja (2004), Manchester and Mumford (2012), and Lindbeck et al. (1999) will also be mentioned in the related context.

Weisbrod (1970) is the very first economic study on welfare stigma and has served as a bridge between earlier sociological studies and emerging economic studies. Weisbrod (1970) is characterised primarily by the idea that welfare stigma is the psychological cost incurred by those participating in a welfare programme[6]. Based on this idea, it argues that the participation decision can be analysed from the cost-benefit perspective.

Weisbrod (1970) is characterised by two distinguishing aspects. First, it differentiates between internal and external types of stigma. Internal type of stigma refers to psychological costs that solely originate from the participants' state of mind and, hence, its intensity is not affected exogenously. On the other hand, external type of stigma refers to psychological costs that incurred for exogenous reasons and, therefore, its intensity is affected by such factors as public exposure and peer effects. In theory, external type of stigma increases with the level of public exposure; that is, the higher the level of public

[5]Although Blumkin et al. (2015) covers across three categories, it places its emphasis on the claim that welfare stigma could potentially be used to improve the efficiency of the targeting based on its analysis of statistical stigma.

[6]In this paper, the phrases *participating in a welfare programme* and *receiving welfare benefits* are used with exactly the same meaning.

exposure, the greater the psychological costs. Meanwhile, it decreases in relation to peer effects; for instance, the closer the geographical proximity between recipients, the smaller the psychological costs[7].

Second, it associates welfare stigma only with psychological costs. Participating in a welfare programme comes with various costs; for example, in addition to the psychological costs that have been discussed so far, participants could face opportunity costs in terms of the time they spend to fill out an application form or visit social welfare offices. Weisbrod (1970) makes a clear distinction between these two types of costs and argues that only psychological costs are classified as welfare stigma.

Manchester and Mumford (2012) separately estimates both opportunity and psychological costs using the data from two food assistance programmes in the United States. More specifically, Manchester and Mumford (2012) interprets variables costs as time costs (opportunity costs) and fixed costs as welfare stigma (psychological costs); then, it estimates each of them based on the generic static model of labour supply, similar to the one proposed by Keane and Moffitt (1998) and finds that welfare stigma is four times as large as time costs. However, due to the difficulties of estimating them separately, most empirical studies do not distinguish the two types of costs from each other. The assumptions in this paper with regard to this issue will be explained in Section 4.

While Weisbrod (1970) is acknowledged as the very first economic study on welfare stigma, it was Moffitt (1983) that provided an econometric framework and conducted an empirical examination of the existence of such a phenomenon for the first time. In contrast to Weisbrod (1970) that defines welfare stigma as psychological costs, Moffitt (1983) defines it as disutility that arises from participating in a welfare programme. The definition by Moffitt (1983) enables the participation decision to be interpreted as the result of an individual's utility maximising behaviour. It follows that Moffitt (1983) introduces a utility function that depends on such variables as the level of disposable income and the actual participation decision, and it explains that an individual will participate in a welfare programme if and only if his or her utility with participation exceeds that without

[7]Needless to say, geographical proximity is not the only cause of peer effects; for example, the rise in the take-up rates could also strengthen such effects and ultimately lead to a lowering of psychological costs.

participation.

Moreover, using the Panel Study of Income Dynamics (PSID), Moffitt (1983) empirically examines whether welfare stigma exists in the Aid to Families with Dependent Children (AFDC) and concludes that the AFDC indeed suffers from welfare stigma. This utility-based approach is applicable to other social welfare programmes; for example, Levinson and Rahardja (2004) applies a similar approach to examine Medicaid and the Food Stamp Program in the Untied States and confirms the existence of welfare stigma.

The significance of Moffitt (1983) also lies in its finding that welfare stigma is only associated with fixed costs. Although its theoretical and empirical models incorporate the possibility that welfare stigma has both flat and variable components, the empirical examination only finds evidence in favour of a flat component and finds no evidence that suggests a variable component.

A flat component implies that an individual suffers a fixed amount of disutility as a result of participation; therefore, an individual is more likely to participate if the amount of welfare benefit increases because the relative level of disutility is decreasing in the amount of welfare benefit. On the other hand, a variable component implies that an amount of disutility is not fixed and participants will receive lower marginal utility from welfare benefit than from private income. It should be noted that Moffitt (1983) only deals with internal type of stigma and, hence, its results do not necessary indicate that external type of stigma also has a flat component alone.

How, then, has the external type of stigma been studied in economics? Besley and Coate (1992) is one of the earliest works that studies this type of welfare stigma and has influenced subsequent economic studies such as Lindbeck et al. (1999). Besley and Coate (1992) develops two independent models based on two different views of external type of stigma, both of which are explained below.

The first model is based on the statistical discrimination view of external type of stigma. This view, which originally appears in Goffman (1963), assumes that there are two types of people among those who receive welfare benefits: on the one hand, there are deserving people, e.g. disabled or elderly people, who are in desperate need of social welfare

owing to non-voluntary circumstances and not because of the lack of 'socially valuable characteristics' (Besley and Coate, 1992); on the other hand, there are undeserving people who need social benefits because of 'blemishes of individual character' (Goffman, 1963) such as laziness or substance abuse.

In general, those who are not in need of public assistance are favourable towards the deserving people but are unfavourable towards the underserving people. This tendency is less problematic when information is symmetrical because the deserving people would not feel discomfort from the non-recipients, assuming that eligibility is the only factor that makes the non-recipients resentful towards the recipients. In reality, however, this is not always the case. Since it is difficult to tell who deserve welfare benefits and who do not, the non-recipients could display a negative attitude not only towards the underserving people but also towards the deserving people. As a result, in this view, all the recipients feel discomfort from the non-recipients, eligibility notwithstanding.

Besley and Coate (1992) argues that this kind of external type of external stigma could be reduced by conducting a better means-test so that only the deserving people would be able to claim welfare benefits. Imposing work requirements could also help to reduce the stigma because those who have an ability to work would no longer be able to deliberately choose not to work and solely live on welfare benefits; in other words, self-targeting is likely to come into effect with work requirements.

The second model is based on the taxpayer resentment view of external type of stigma. This is caused by the fact that welfare benefits are covered by tax payments. Besley and Coate (1992) argues that some, if not all, taxpayers[8] would feel dissatisfied that the taxes they pay are not used for themselves but for those who are unemployed or who have low income. Therefore, the non-recipients could negatively perceive not only the undeserving people but also the deserving people since whether he or she is deserving does not change the fact that part of their tax payments are used for welfare benefits. In this case, the level of welfare stigma is increasing in the amount of welfare benefits, ceteris paribus, because taxpayers are likely to grow a stronger feeling that the participants are receiving too much.

[8] Although in reality the recipients with low income also pay taxes, for the sake of simplicity it is assumed here that taxpayers do not participate in a welfare programme.

An economic model is derived from each of these two views and each model can be solved independent for equilibrium independently. Interestingly, while it assumes unique equilibrium for the purposes of conducting a comparative static analysis, Besley and Coate (1992) mentions the possibility that equilibrium may not be unique. This implication has opened up a way for a dynamic analysis of external stigma; for instance, in its study of the impact of social norm on economic and political decisions, Lindbeck et al. (1999) incorporates the idea of welfare stigma having multiple equilibria. Further research on the dynamic aspect of welfare stigma, especially the empirical examination of the existence of multiple equilibria, is highly valuable as the results could reveal new attributes of welfare stigma.

Moffitt (1983) and Besley and Coate (1992) can be viewed as favouring the minimisation of welfare stigma based on the idea that welfare stigma is preventing those who deserve welfare benefits from applying for them and is causing economic inefficiency; in other words, these studies deal with traditional stigma. However, as discussed in the previous sub-section, statistical stigma could reduce incentives to falsely receive welfare benefits. In fact, Blumkin et al. (2015) claims that employing statistical stigma as a policy instrument could lead to the improved targeting and higher efficiency of welfare programs; therefore, it questions the view that welfare stigma should be unconditionally reduced.

There have been many economic studies on welfare stigma in the United States and European countries. To the best of author's knowledge, however, there have been no such studies in Japan. Unfortunately, compared to the number of sociological or psychological studies, there has been, more generally, a much smaller number of economic studies on public assistance carried out in Japan in the first place. The shortage of relevant data available is likely to be the biggest obstacle for conducting such studies. Since large scale surveys such as the PSID in the United States did not exist in Japan for quite a long time, most of the past economic studies on public assistance in Japan had no choice but to use macro data on the administrative divisions level. However, participation decisions cannot be analysed without micro data on the individual level. In this respect, the JHPS/KHPS is useful in the empirical analysis of this paper in that it provides the micro data on

thousands of Japanese citizens and that it can therefore be applied to the spatial probit models to be proposed in Section 4.

3 Public Assistance System in Japan

The public assistance system, or *seikatsu hogo seido*, in Japan was established under the Public Assistance Act of 1950. The fundamental objective of the public assistance system is to guarantee a minimum standard of living, according to Article 1 of the New Act[9], and this objective complements Article 25 of the Constitution of Japan which states that 'All people shall have the right to maintain the minimum standards of wholesome and cultured living'.

The public assistance system in Japan is based on three principles: Nondiscrimination and Equality (Article 2); Minimum of Standard of Living (Article 3); and Supplementary Nature of Public Assistance (Article 4). Out of these principles, all of which reflect Article 1 of the New Act, the third principle is worth further attention.

In general, under the social security system, the public assistance system is responsible for relieving those who fall into poverty even with social insurance; in other words, the public assistance system plays a role as the last resort. In this respect, the third principle, which requires recipients to fully utilise every available resource, indicates that the public assistance system in Japan is no exception to this general idea and indeed serves a role as the last resort.

Not surprisingly, the third principle is also reflected in the method of the calculation. The calculation method is briefly explained below, based on the Ministry of Health, Labour and Welfare (2015) and Hayashi (2010). The amount of public assistance equals the difference between the minimum cost of living and the amount of income appropriation. The minimum cost of living is the sum of the minimum costs calculated for each of the eight categories: (i) livelihood, (ii) education, (iii) housing, (iv) medical care, (v) long-term care, (vi) childbirth, (vii) employment, and (viii) funerals and other ceremonies. Note that

[9]The Public Assistance Act was originally enacted in 1946 but revised in 1950; therefore, the Public Assistance of 1950 is often called the *New* Public Assistance Act. In this paper, the 1950 Act is henceforth referred to as the New Act.

the standard for livelihood assistance takes into account household characteristics, such as household size, household composition, and area of residence. For example, Table 3 shows the livelihood assistance standards for four types of households in 2015. According to Table 3, the standard for livelihood assistance for a three-person household (male aged 33, female aged 29, and child aged 4) living in the first class area – 1 is 160,110 Japanese yen (henceforth, JPY).

See Figure 1 for the trend in the public assistance rate since 1951. It clearly shows the sharp decline in the public assistance rate during the two periods (from the 1950s to the 1960s and from the late 1980s to the mid-1990s). The first decline was the result of such events as the rapid growth of the Japanese economy and the expansion of the social insurance system, which led to the achievement of universal health insurance coverage in 1961. The public assistance rate was stable around 12 per mil in the 1970s and the early 1980s, yet it started to drop once again. The second decline occurred from the late 1980s to the mid-1990s, by and large coinciding with the economic bubble in Japan; before long, the public assistance rate hit the lowest level of seven per mil in 1995.

However, the collapse of the bubble economy in the early 1990s, which marked the beginning of the *lost decades*, badly influenced the employment situation. In the aftermath, the landscape of the social security system shifted and the downward trend in the public assistance rate reversed. As shown in Figure 1, the rate has been on the rise since 1995 and hit 17 per mil in 2014. Such a rise in the public assistance rate after the mid-1990s can be understood from the perspective of the social security system, which is explained right below.

In Japan, it is argued that the social insurance system and the Japanese Employment System are interrelated, the latter being defined as a mutually complementary relationship among various employment systems established during the late 1960s and the early 1970s (Saguchi, 2015). Due to this mechanism, the stagnation of the Japanese Employment System, along with fiscal difficulties, after the collapse of the bubble economy posed a challenge to the social insurance system. As the social insurance system became unstable and its sustainability began to be called into question, it was public assistance that quickly

gained importance within the social security system. Just like the case in the United States, however, the issue of welfare stigma had kept the public assistance system in Japan underdeveloped throughout the 20$^{\text{th}}$ century.

In light of the fact that the public assistance rate in Japan has been rising for the last two decades, it is considered to be more urgent than ever to improve the public assistance system so that it can play a central role in the *new normal* of the Japanese social security system, in which the social insurance system is not as significant as it was in the 20$^{\text{th}}$ century. Taking the current situation into account, this paper will provide empirical evidence with regard to the existence of welfare stigma in Japan.

As clarified in Section 1, this paper will only focus on internal and external types of traditional stigma; in other words, it is assumed that statistical stigma can be ignored. There are two reasons for this assumption. First, having considered that a theoretical model proposed by Blumkin et al. (2015) could be considered to incorporate both internal and external types of statistical stigma is already proposed by Blumkin et al. (2015), this paper started as an attempt to fill the gap in the existing literature by proposing a model that incorporates internal and external types of traditional stigma. Second, contrary to popular belief that has been fuelled by the media, the benefit fraud rate is actually low in Japan and, hence, statistical stigma is not considered to be as problematic as traditional stigma; according to the Ministry of Internal Affairs and Communications (2014), the benefit fraud rate is 1.4 per cent in 2010, 1.6 per cent in 2011, and 1.6 per cent in 2012.

Based on these reasons, it is assumed here that statistical stigma can be ignored for the purposes of this paper; therefore, the empirical analysis in this paper only concerns traditional stigma. It follows that, for the sake of simplicity, traditional stigma will from now on simply be referred to as stigma unless otherwise specified. Needless to say, however, statistical stigma would be worthy of further investigation if benefit fraud becomes more serious in the future; in addition, there is also plenty of room to develop a model that incorporates both traditional and statistical stigmas.

4 Model

The spatial autoregressive probit model (henceforth, SAPM) and the spatial error probit model (henceforth, SEPM) are used to empirically examine the existence of welfare stigma in Japan. This section consists of three parts: the first part introduces the SAPM; the second part introduces the SEPM; and the third part outlines some assumptions. For mathematical details, see also LeSage and Pace (2009).

Suppose that an individual i $(= 1, ..., n)$ participates in a public assistance programme if and only if $P_i^* > 0$, where P_i^* is an unobservable latent variable that indicates predisposition towards participation. Formally:

$$P_i = \begin{cases} 1 & (P_i^* > 0) \\ 0 & (P_i^* \leq 0) \end{cases} \tag{1}$$

where P_i represents i's participation status. Eq. (1) is assumed to apply to both models.

To begin with, the SAPM specifies P_i^* as

$$\begin{gathered} P_i^* = \alpha + \beta' X_i + \rho \sum_{j=1}^{n} w_{ij} P_j^* + \varepsilon_i \\ \varepsilon_i \sim N(0, \sigma^2 I_n) \end{gathered} \tag{2}$$

where α is the constant term parameter, β' is a vector of coefficients associated with a vector of independent variables X_i, ρ is a scalar parameter, and w_{ij} is a weight that takes the value of one if individuals i and j are neighbours. Using a matrix notation, Eq. (2) can be re-written in the following reduced form:

$$P^* = (I_n - \rho W)^{-1}(\alpha \iota_n + X\beta + \varepsilon) \tag{3}$$

where I_n is an n-dimensional identity matrix and ι_n is an $n \times 1$ vector of ones associated with α. Also, W is an $n \times n$ spatial weight matrix and its general form is written as

follows:

$$W = \begin{pmatrix} 0 & w_{12} & \dots & w_{1n} \\ w_{21} & 0 & \dots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \dots & 0 \end{pmatrix} \tag{4}$$

The specification of W in this empirical analysis will be explained in the next section.

Eq. (2) involves two coefficients that represent the existence of welfare stigma. First, α represents fixed costs associated with public assistance and, hence, it is an indicator of internal type of stigma. More specifically, internal type of stigma is thought to exist if and only if α is significantly negative. Second, ρ is an indicator of external type of stigma, which in this case is caused by geographical proximity among recipients, and there is the possibility that external stigma exists if and only if ρ is significantly positive. At this point, however, its existence cannot be confirmed because this potential strategic behaviour could be the result of spatially correlated shocks. Therefore, in a case where ρ in Eq. (2) is significantly positive, it becomes all the more necessary to estimate the SEPM so that it can be clarified whether the potential strategic behaviour reflects spatially correlated shocks.

Next, the SEPM specifies P_i^* as

$$\begin{aligned} P_i^* &= \alpha + \beta' X_i + u_i \\ u_i &= \lambda \sum_{j=1}^{n} w_{ij}\mu_i + \varepsilon_i \\ \varepsilon_i &\sim N(0, \sigma^2 I_n) \end{aligned} \tag{5}$$

where λ is a scalar parameter. In the same manner as the SAPM, Eq. (5) can be re-written in the following reduced form:

$$P^* = (\alpha \iota_n + X\beta) + (I_n - \lambda W)^{-1}\varepsilon \tag{6}$$

The SEPM is estimated to investigate the presence of spatially correlated shocks, which is

indicated by the significant estimate of λ. Therefore, it can be concluded that the external type of stigma exists if and only if the results satisfy the following two conditions at the same time: first, ρ is significantly positive; and second, (ii) λ is not significant.

There are various ways to estimate the spatial models. For instance, McMillen (1992) proposes maximum likelihood (ML) estimation using the expectation-maximisation (EM) algorithm. Also, Pinkse and Slade (1998) proposes generalised method of moments (GMM) estimation and applies it to a spatial logit model. McMillen's approach, however, has some drawbacks; for example, its use of a covariance matrix conditional on ρ causes a downward bias on confidence intervals (LeSage, 2000). In an endeavour to overcome such drawbacks, LeSage (2000) proposes Bayesian Markov Chain Monte Carlo (MCMC) estimation, extending a prior study by Albert and Chib (1993). Compared to the ML and GMM approaches, both of which are considered to be associated with computational burden (Arbia, 2014), the Bayesian MCMC approach is considered to be computationally more efficient than the ML or GMM approach[10]. Due to this attribute, this approach has been used to estimate the spatial probit models in various studies (e.g. Kakamu et al. (2010), Jaimovich (2012), Ghosh (2013), Loomis and Mueller (2013), and Nikolic and Weiss (2014)). The empirical analysis of this paper also relies on Bayesian inference based on Gibbs sampling, an MCMC method, to estimate the SAPM and the SEPM. Technical details on the implementation of Bayesian MCMC estimation will be addressed in the next section.

Finally, last but not least, the above models are based on the following three assumptions. The first assumption is that internal type of stigma does not have a variable component; in other words, internal type of stigma is assumed to be fixed costs that arise from participation. This assumption is based on the following idea. Internal type of stigma having a variable component means that participants receive lower marginal utility from welfare benefit than from private income (Moffitt, 1983). Unlike the Food Stamp Program in the United States, however, public assistance in Japan is an in-cash benefit; thus, as far

[10]Klier and McMillen (2008) proposes linearized GMM estimation, which is computationally more efficient than (non-linearized) GMM estimation. To the author's knowledge, however, it has only been applied to a spatial logit model; hence, even though a function `spprobit` in an `R` package `McSpatial` supports linearized GMM estimation of the SAPM, this approach is not adopted in this empirical analysis.

as the public assistance programme in Japan is concerned, the marginal utility must be equal between cash from public assistance and cash from private income. In fact, providing support for this assumption, Moffitt (1983) finds no evidence of a variable component in the AFDC, which is also an in-cash benefit.

The second assumption is that any costs from participation are associated with welfare stigma. In reality, costs from participation consist of time costs and physical costs, and only physical costs are classified as welfare stigma (Weisbrod, 1970). This assumption is imposed for two reasons: first, since this paper aims to examine whether welfare stigma exists in Japan, separately estimating time and physical costs is another issue that is left for future analysis; second, separating the two types of costs is technically difficult due to the limitations of the data. Still, this assumption makes a credible starting point considering that Manchester and Mumford (2012) reports psychological costs comprise 80 per cent and time costs 20 per cent of total costs from participation in food assistance programmes in the United States.

The third assumption is that external type of stigma is solely caused by peer effects based on geographical proximity. In theory, there could be multiple causes for external type of stigma; for example, whereas peer effects encourage those eligible to participate in the public assistance programme, the fear of public exposure discourages them from doing so. The latter aspect of external type of stigma is hard to quantify in the first place and, as far as the empirical analysis of this paper is concerned, the limitations of data make it all the more difficult. Thus, for the time being, external type of stigma is assumed to be solely caused by peer effects.

5 Data and Methodology

5.1 Data

The empirical analysis of this paper primarily uses the 2015 wave[11] of the Japan Household Panel Survey (JHPS/KHPS) and partially relies on the 2009 wave for additional

[11]The 2015 wave was conducted on 31st January, 2015.

information on the education levels of the heads of households. It is a nationwide household panel survey that has been conducted by the Panel Data Research Center at Keio University since 2009. The JHPS/KHPS provides large data sets regarding respondents' socio-economic status and has prompted empirical research on social policies in Japan. In the context of this paper, the JHPS/KHPS has an advantage over other surveys in Japan, e.g. the Japanese General Social Surveys (JGSS), in that it provides the actual amount of private income, public assistance, and other welfare benefits.

See Table 4 for an overview of the respondents. In Table 4, the vertical and horizontal axes represent the marital and participation statuses of respondents, respectively. Out of 2,198 respondents in the original data set, 583 married and 158 unmarried respondents who do not reveal their participation status are excluded from the sample of this empirical analysis. In addition, eight respondents, all of whom happen to be non-participants, are excluded from the sample since the education level of the heads of their households cannot be identified. The details on the exclusion of these eight respondents are given in the next sub-section.

In the end, the sample size is 1,450. According to Table 4, five out of 1,073 married respondents and 15 out of 377 unmarried respondents received public assistance in 2015. It follows that the public assistance rate within this sample is 13.8 per mil (20/1,445). This rate is lower than 17 per mil, the population-based public assistance rate in 2014, and 31.6 per mil, the household-based public assistance rate in 2013 (Ministry of Health, Labour and Welfare, 2014, 2016a)[12].

Needless to say, it would be desirable if the participation rate within the sample matches the actual rate. Just like the case above, however, the two rates do not always coincide; for example, Moffitt (1983) also reports a similar issue and argues that under-reporting may be part of the cause. Considering that such comprehensive surveys as the JHPS/KHPS were not common in Japan until recently, under-reporting of participation can also be presumed to be the case here. It should be noted, however, that under-reporting itself could be the consequence of welfare stigma. In this respect, empir-

[12]The population-based and household-based public assistance rates are the percentages of Japanese citizens and households that receive public assistance, respectively.

ical studies of welfare stigma are in a dilemma: on the one hand, welfare stigma cannot be precisely examined unless all respondents reveal their true participation status, but on the other hand, there will be much less to be examined since welfare stigma is almost, if not entirely, non-existent in such a circumstance.

As for the dependent variable, this empirical analysis uses a binary variable PAR that represents participation status. PAR is defined based on the amount of public assistance received by the household (PUB; one unit = 10,000 JPY) as follows:

$$\mathrm{PAR}_i = \begin{cases} 1 & (\mathrm{PUB}_i > 0) \\ 0 & (\mathrm{PUB}_i = 0) \end{cases} \tag{7}$$

See Table 5 for the list of independent variables. A vector X_i in Eq. (2) and (5) contains the following independent variables: age of the head of household (AGE); marital status of the head of household (MAR); dummy variable for female-headed households (FEM); education level of the head of household (EDU); employment status of the head of household (EMP); number of children of 18 or under (CLD); amount of private income received by the household (PRI); and amount of other welfare benefits[13] received by the household (WEL). Moreover, a variable for region of residence (REG) is used in the specification of a weight w_{ij} in Eq. (2) and (5); the specification of w_{ij} is explained in the next sub-section. See also Table 6 for descriptive statistics.

5.2 Methodology

This sub-section consists of three parts: the first part explains the treatment of data; the second part explains the specification of a weight matrix W; and the third part addresses technical details of Bayesian MCMC estimation.

To begin with, several notes regarding the treatment of data are addressed below. First, the age of the head of household is calculated as of 31$^{\mathrm{st}}$ January, 2015, using year and month of birth of each household head. For the sake of confidentiality, day of birth of respondents and their spouses is omitted from the original data set; furthermore, that

[13] Other welfare benefits include unemployment benefits, childcare leave benefit, and childcare allowance.

of other members of the families is not even included in the questionnaire.

Second, as for marital status, if the head of household is a parent of the respondent and both parents are still alive, then it is assumed that the parents are married, i.e. MAR = 1. Indeed, an increase in the number of late-life divorces, or – in common Japanese parlance – *jukunen rikon*, in Japan has often been reported on the news for quite some time. For the sake of simplicity, however, this empirical analysis is based on the premise that there are more married elderly couples than divorced elderly ex-couples.

Third, education level is based on the 2009 wave and not on the 2015 wave of the JHPS/KHPS. This is because education level is comprehensively surveyed in the 2009 wave, and only the changes since then are reported in the 2015 wave. Although there is a six-year interval between the two waves, referring to the 2009 wave for education level would not be a serious issue. All 1,450 household heads are over 22 years of age as of 31st January, 2009; therefore, even though it is possible that some of the household heads began pursuing higher education after 2009, such changes are unlikely to cause a significant impact on the overall results.

Fourth, based on the 2009 wave, the education level of each head of household is specified according to the flow diagram described in Figure 2. By means of this procedure, two issues with the data can be tackled: the first issue is that, due to the limitations of data, education level of the head of household cannot be identified if neither a respondent nor his/her spouse is the household head[14]; the second issue is that, even when either respondents or their spouses are the heads of households, data on education level is missing in some cases.

As mentioned briefly in the previous sub-section, listwise deletion is performed on eight household heads whose education level cannot be specified. The method of listwise deletion is used on these samples based on the following evidence. See Table 7 for the results of the Little's MCAR test that is performed on all variables used in this empirical analysis. The Little's MCAR test is a test to examine whether the choice of deletion targets is independent of the data values, i.e. data is missing completely at random (MCAR)[15].

[14] 33 out of 1,450 samples fall under this case.
[15] Refer to Little (1988) for further details.

According to Table 7, the null hypothesis that deletion targets are chosen completely at random is not significant at the five per cent level of significance; therefore, listwise deletion can be considered to be an effective approach to the missing data and parameter estimates will not be biased as a result of this method (Allison, 2010).

Fifth, last but not least, the amount of private income is calculated according to the following two rules. The first rule is that, if the total income exceeds the sum of the sources of income by more than 500,000 JPY and there is at least one non-response in the sources of income, then the amount of private income (PRI) is calculated as the difference between the total income and the amount of welfare benefits[16]. The second rule is that, if the total income of respondents is not available due to non-responses, then the amount of private income is calculated as the difference between the sum of the sources of income and the amount of welfare benefits. The calculation relies on these rules because the JHPS/KHPS asks respondents their total income and sources of income in separate questions and, as a result, there is a significant difference between the two values, which should be equal by definition, for quite a few respondents.

Next, the specification of a weight matrix W is explained. In this empirical analysis, W, a $1,450 \times 1,450$ square matrix, is necessary to investigate the existence of peer effects among participants in the public assistance programme, since ρ in Eq. (2) and λ in Eq. (4) cannot be estimated without W being obtained. See Eq. (4) for the general form of W. Now, an element w_{ij}, the (i, j) entry of W, is defined as follows:

$$w_{ij} = \frac{W_{ij}}{\sum_{j=1}^{n} W_{ij}}$$

$$\text{where } W_{ij} = \begin{cases} 1 & \text{if individuals } i \text{ and } j \text{ are neighbours} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

In this empirical analysis, two individuals are defined as *neighbours* if they live in either the same or neighbouring regions. See Table 8 for the contiguous relationship among eight regions of Japan, in which the value of one indicates that two regions are next to

[16]The amount of welfare benefits, which is referred to here, is the sum of the amount of public assistance (PUB) and that of other welfare benefits (WEL).

each other. For instance, if individuals i and j live in the Tohoku and Kanto regions, respectively, then w_{ij} equals one. In terms of technical details, W can be specified in `R` by a function `mat2listw`, which is available in a package `spdep`.

Finally, an `R` package `spatialprobit` is used to perform Bayesian MCMC estimation. More specifically, Bayesian MCMC estimation of the SAPM and the SEPM is implemented with the functions `sarprobit` and `semprobit`, respectively.

6 Results

This section presents the results and analysis of empirical examinations in three parts: the first part presents the estimation results of the standard probit model, which serve as the benchmark for other parts; the second part presents those of the SAPM; and the third part presents those of the SEPM.

First of all, column (1) of Table 9 shows the Bayesian MCMC estimation results of the standard probit model, which does not take account of potential spatial effects. This estimation can be implemented by a function `MCMCprobit` in an `R` package `MCMCpack`. According to the results, an indicator of internal type of stigma, α, is negative but not significant at any level; therefore, Bayesian MCMC estimation of the standard probit model does not reject the null hypothesis that internal type of stigma is absent. The estimation finds that the one associated with private income is the only significant coefficient estimate; that is, the negatively significant coefficient of private income indicates that the likelihood of participation is decreasing in its amount. Although the coefficient estimates of other independent variables are not significant at any level, the signs of some variables are consistent with the theory. For example, the results suggest that female-headed households are more likely to receive public assistance than male-headed households; that having more children raises the likelihood of being recipients; and that an increase in welfare benefits other than public assistance alleviates the need for public assistance. On the other hand, despite none of them being significant at any level, the signs of four out of five dummy variables on education levels suggest that completing higher education actually increases

the likelihood of a household receiving public assistance, which is not consistent with the common notion.

Next, column (2) of Table 9 shows the estimation results of the SAPM. In contrast to the previous case, α is estimated to be significantly negative at -3.4162, which indicates that public assistance is associated with fixed costs; therefore, Bayesian MCMC estimation of the SAPM reveals that internal type of stigma exists. Meanwhile, ρ is significantly negative at -0.0034. Recall that the existence of external type of stigma is confirmed if and only if ρ is significantly positive, as this paper assumes it to be only caused by peer effects which positively affect take-up. Therefore, the results suggest that external type of stigma is absent. Moreover, the coefficient estimates of dummy variables on education levels are now all negative, yet still not significant; the coefficient estimates of other variables except for that of age have the same signs as the previous case. Thus, as far as the signs of the coefficient estimates are concerned, the estimation results of the SAPM seem to be more consistent with the theory than those of the standard probit model.

Finally, column (3) of Table 9 shows the estimation results of the SEPM. Since the spatial autoregressive probit estimation suggests the absence of external type of stigma, estimating the SEPM in this circumstance is not so significant as doing so under its existence. Having said that, it is still worthwhile to estimate the SEPM so that the presence of spatially correlated shocks can be empirically examined. According to the results, the coefficient of λ, an indicator of such shocks, is estimated to be 0.2767 but is not significant at any level; therefore, the results suggest that the spatially correlated shocks are not present, further confirming the absence of any spatial effect whatsoever. It should be noted, however, α is now estimated to be neither negative nor significant.

7 Conclusion

Despite the high public attention, issues concerning the current public assistance system in Japan have rarely been studied in the field of economics. One such issue is welfare stigma, which has been regarded as a cause of a low take-up of public assistance. Consequently,

this paper has focused on this particular issue and has empirically examined its existence in Japan. More specifically, it has performed an empirical analysis by estimating the SAPM and the SEPM based on the JHPS/KHPS. The application of spatial econometrics in the analysis of welfare stigma has an advantage that internal and external types of traditional stigma can be examined at the same time.

The estimation results of the spatial probit models do not reject the existence of internal type of traditional stigma but reject the existence of external type of traditional stigma. More specifically, the estimation results of the SAPM show that α is significantly negative at -3.4162, suggesting the strong existence of internal type of traditional stigma. However, ρ is estimated to be not only significant but also negative at -0.0034; hence, external type of traditional stigma does not exist. Moreover, Bayesian MCMC estimation of the SEPM finds that λ is not significant at any level, suggesting that the spatially correlated shocks do not exist, either. In summary, the empirical analysis of this paper has revealed that internal type of traditional stigma indeed exists but that external type of traditional stigma, or any spatial effect whatsoever, does not.

This paper contributes to the existing literature in the following two ways. First, it is the first study to empirically examine the existence of welfare stigma in Japan and finds that neither internal nor external type of traditional stigma exists in Japan. Second, it proposes the application of the spatial probit models to the analysis of welfare stigma for the first time and demonstrates that internal and external types of traditional stigma can be simultaneously estimated.

As already mentioned, the estimation results suggest that welfare stigma does not exist in Japan. In other words, according to the results, there are no psychological barriers that prevent those eligible from participating in the public assistance programme. Having said that, it is important to acknowledge that the results may not always be the case since this empirical analysis is based on a number of assumptions and the limited data available.

For example, one of the extreme assumptions lies in the definition of *neighbours*. In the empirical analysis of this paper, whether two samples are *neighbours* is judged based on their regions of residence; needless to say, however, the idea of contiguity is too broadly

interpreted under this definition, making the assumption of strategic behaviour hardly realistic. In addition, low take-up of public assistance is another issue that needs to be resolved (Fujisawa, 2008). That is, the absence of welfare stigma implies that there are other factors that somehow prevent those eligible from applying for public assistance. In accordance with the aim of this paper, however, such factors are not investigated.

Hence, the empirical analysis of this paper leaves plenty of room for future research. Some avenues for further development are suggested below. First, this paper identifies external type of traditional stigma solely based on the peer effects among recipients, so it would be valuable to take account of public exposure that will enhance the stigma.

Second, since the models presented in this paper are applicable to other survey data, it would be worthwhile to estimate these models using another survey that puts emphasis on low-income households, e.g. the National Survey of America's Families (NSAF), so that welfare stigma is more accurately identified.

Third, although this paper assumes that statistical stigma can be ignored, there has been a lack of empirical studies on statistical stigma in the United States or European countries, let alone in Japan; hence, it would be of interest to investigate whether this assumption is credible.

Fourth, despite a further attempt to compare the ML and Bayesian MCMC estimation results, ML estimation of the above-specified spatial probit models was not completed in this empirical analysis due to the technical difficulties faced by the author. Although an `R` function `spprobitml` in a package `McSpatial` can implement ML estimation of the SAPM[17], its underlying model does not include an intercept α. For the purposes of this empirical analysis, the source code needs to be modified in such a way that it includes α; however, the author is incapable of such a modification as yet. Moreover, an `R` function `SpatialProbitFit` in a package `ProbitSpatial` can implement ML estimation of the SAPM and the SEPM, yet the calculation of the standard errors of ρ and λ has been unsuccessful and the author has yet to identify its cause. The resolution of these technical difficulties is also left for future work.

[17]Note that this package does not provide a function for ML estimation of the SEPM in the first place.

References

Albert, J. H. and Chib, S. (1993). Bayesian Analysis of Binary and Polychotomous Response Data. *Journal of the American Statistical Association*, 88(422):669–679.

Allison, P. D. (2010). Missing data. In Wright, J. D. and Marsden, P. V., editors, *Handbook of Survey Research*, pages 631–657. Emerald Group Publishing Ltd.

Arbia, G. (2014). *A Primer for Spatial Econometrics*. Palgrave Macmillan.

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

Blumkin, T., Margalioth, Y., and Sadka, E. (2015). Welfare Stigma Re-examined. *Journal of Public Economic Theory*, 17(6):874–886.

Fujisawa, M. (2008). Poverty and Social Assistance System in Japan: Analysis of Take-up Rate Behavior of Social Assistance based on The JGSS Data [In Japanese]. JGSS Research Series No. 4.

Ghosh, S. (2013). Participation in school choice : a spatial probit analysis of neighborhood influence. *Annals of Regional Science*, 50:295–313.

Goffman, E. (1963). *Stigma: Notes on the Management of Spoiled Identity*. Prentice-Hall, New Jersey.

Hayashi, M. (2010). Social Protection in Japan: Current State and Challenges. In Asher, M. G., Oum, S., and Parulian, F., editors, *Social Protection in East Asia - Current State and Challenges*, pages 19–54. ERIA Research Project Report 2009-9, Jakarta: ERIA.

Jaimovich, D. (2012). A Bayesian spatial probit estimation of Free Trade Agreement contagion. *Applied Economics Letters*, 19(6):579–583.

Kakamu, K., Wago, H., and Tanizaki, H. (2010). Estimation of Regional Business Cycle in Japan using Bayesian Panel Spatial Autoregressive Probit Model. In Nolin, T. P., editor, *Handbook of Survey Research*, pages 555–571. Nova Science Publishers, Inc., New York.

Kayser, H. and Frick, J. R. (2000). Take It or Leave It: (Non-)Take-Up Behavior of Social Assistance in Germany. Discussion Paper No. 210. DIW Berlin.

Keane, M. and Moffitt, R. (1998). A Structural Model of Multiple Welfare Program Participation and Labor Supply. *International Economic Review*, 39(3):553–589.

Klier, T. and McMillen, D. P. (2008). Clustering of Auto Supplier Plants in the United States: Generalized Method of Moments Spatial Logit for Large Samples. *Journal of Business & Economic Statistics*, 26(4):460–471.

LeSage, J. and Pace, R. K. (2009). *Introduction to Spatial Econometrics*. CRC Press.

LeSage, J. P. (2000). Bayesian Estimation of Limited Dependent Variable Spatial Autoregressive Models. *Geographical Analysis*, 32(1):19–35.

Levinson, A. and Rahardja, S. (2004). Medicaid Stigma. Discussion Paper.

Lindbeck, A., Nyberg, S., and Weibull, J. W. (1999). Social Norms and Economic Incentives in the Welfare State. *Quarterly Journal of Economics*, 114(1):1–35.

Little, R. J. A. (1988). A Test of Missing Completely at Random for Multivariate Data with Missing Values. *Journal of the American Statistical Association*, 83(404):1198–1202.

Locke, J. (1791). *An Account of the Origin, Proceedings, and Intentions of the Society for the Promotion of Industry in the Southern District of the Parts of Lindsey, in the Country of Lincoln*. R. Sheardown, Louth.

Loomis, J. B. and Mueller, J. M. (2013). A Spatial Probit Modeling Approach to Account for Spatial Spillover Effects in Dichotomous Choice Contingent Valuation Surveys. *Journal of Agricultural and Applied Economics*, 45(1):53–63.

Manchester, C. F. and Mumford, K. J. (2012). How Costly is Welfare Stigma? Sperating Psychological Costs from Time Costs. Working Paper.

Martin, P. P. and Weaver, D. A. (2005). Social Security: A Program and Policy History. *Social Security Bulletin*, 66(1):1–15.

McMillen, D. P. (1992). Probit with Spatial Autocorrelation. *Journal of Regional Science*, 32(3):335–348.

Ministry of Health, Labour and Welfare (2000). Annual Reports on Health and Welfare 1998-1999: Social Security and National Life. Available at: `http://www.mhlw.go.jp/english/wp/wp-hw/index.html` (2016/10/20).

Ministry of Health, Labour and Welfare (2010). Report on Social Welfare Administration and Services (FY2009) [In Japanese]. Available at: `http://www.mhlw.go.jp/toukei/saikin/hw/gyousei/09/dl/data.pdf` (2016/11/06).

Ministry of Health, Labour and Welfare (2014). Survey on Public Assistance Recipients (FY2012) [In Japanese]. Available at: `http://www.mhlw.go.jp/toukei/saikin/hw/hihogosya/m2013/dl/kakutei.pdf` (2016/11/06).

Ministry of Health, Labour and Welfare (2015). Annual Health, Labour and Welfare Report 2015, No. 8: Social Welfare and Relief for War Victims. Available at: `http://www.mhlw.go.jp/english/wp/wp-hw9/dl/08e.pdf` (2016/10/20).

Ministry of Health, Labour and Welfare (2016a). Handbook of Health and Welfare Statistics 2015 [In Japanese]. Available at: `http://www.mhlw.go.jp/toukei/youran/data27k/3-03.xls` (2016/11/06).

Ministry of Health, Labour and Welfare (2016b). Survey on Public Assistance Recipients (FY2014) [In Japanese]. Available at: `http://www.mhlw.go.jp/toukei/saikin/hw/hihogosya/m2015/dl/kakutei.pdf` (2016/11/02).

Ministry of Internal Affairs and Communications (2012). Historical Statistics of Japan. Available at: `http://www.stat.go.jp/data/chouki/zuhyou/23-37.xls` (2016/11/06).

Ministry of Internal Affairs and Communications (2014). Fact-Finding Survey on Public Assistance [In Japanese]. Available at: `http://www.soumu.go.jp/main_content/000305409.pdf` (2016/11/03).

Moffitt, R. (1983). An Economic Model of Welfare Stigma. *American Economic Review*, 73(5):1023–1035.

Nikolic, A. and Weiss, C. (2014). Spatial Interactions in Location Decisions: Empirical Evidence from a Bayesian Spatial Probit Model. WU (Vienna University of Economics and Business) Economics Working Paper No. 177.

Pinkse, J. and Slade, M. E. (1998). Contracting in space: An application of spatial statistics to discrete-choice models. *Journal of Econometrics*, 85:125–154.

Saguchi, K. (2015). Japanese Employment System and Industrial Relations: Theory of the Post-War History [In Japanese]. In *The Formation and Development of the 'Japanese Employment System'*, pages 1–69. JTUS Research Institute for Advancement of Living Standards (RENGO-RIALS). Available at: `http://rengo-soken.or.jp/report_db/file/1430367361_a.pdf` (2016/11/03).

Spicker, P. (2011). Stigma and social welfare. Available at: `http://www.spicker.uk/books/Paul%20Spicker%20-%20Stigma%20and%20Social%20Welfare.pdf` (2016/10/20).

Weisbrod, B. A. (1970). On the Stigma Effect and the Demand for Welfare Programs: A Theoretical Note. Discussion Paper 82-70. Institute for Research on Poverty, University of Wisconsin-Madison.

A Figures and Tables

Table 1: Four Categories of Welfare Stigma

	Traditional stigma	Statistical stigma
Internal type	Category I	Category III
External type	Category II	Category IV

Table 2: Representative Economic Literature on Welfare Stigma

Category	Representative literature
I	Moffitt (1983)
	Levinson and Rahardja (2004)
	Manchester and Mumford (2012)
	Blumkin et al. (2015)
II	Besley and Coate (1992)
	Lindbeck et al. (1999)
III	Blumkin et al. (2015)
IV	Blumkin et al. (2015)

Table 3: Examples of Livelihood Assistance Standards in 2015 (Unit: JPY)

	3-person household: male aged 33, female aged 29, child aged 4	Elderly single household: female aged 68
1st class area - 1	160,110	80,870
1st class area - 2	153,760	77,450
2nd class area - 1	146,730	73,190
2nd class area - 2	142,730	71,530
3rd class area - 1	136,910	68,390
3rd class area - 2	131,640	65,560
	Elderly couple household: male aged 68, female aged 65	Single-parent household: female aged 30, children aged 4 and 2
1st class area - 1	120,730	189,870
1st class area - 2	115,620	183,940
2nd class area - 1	109,250	174,860
2nd class area - 2	106,770	171,940
3rd class area - 1	102,090	164,820
3rd class area - 2	97,860	159,900

Source: Ministry of Health, Labour and Welfare (2015)

Table 4: Marital and Participation Statuses of Respondents

	Participants	Non-participants	Total
Married	5	1,068	1,073
Unmarried	15	362	377
Total	20	1,430	1,450

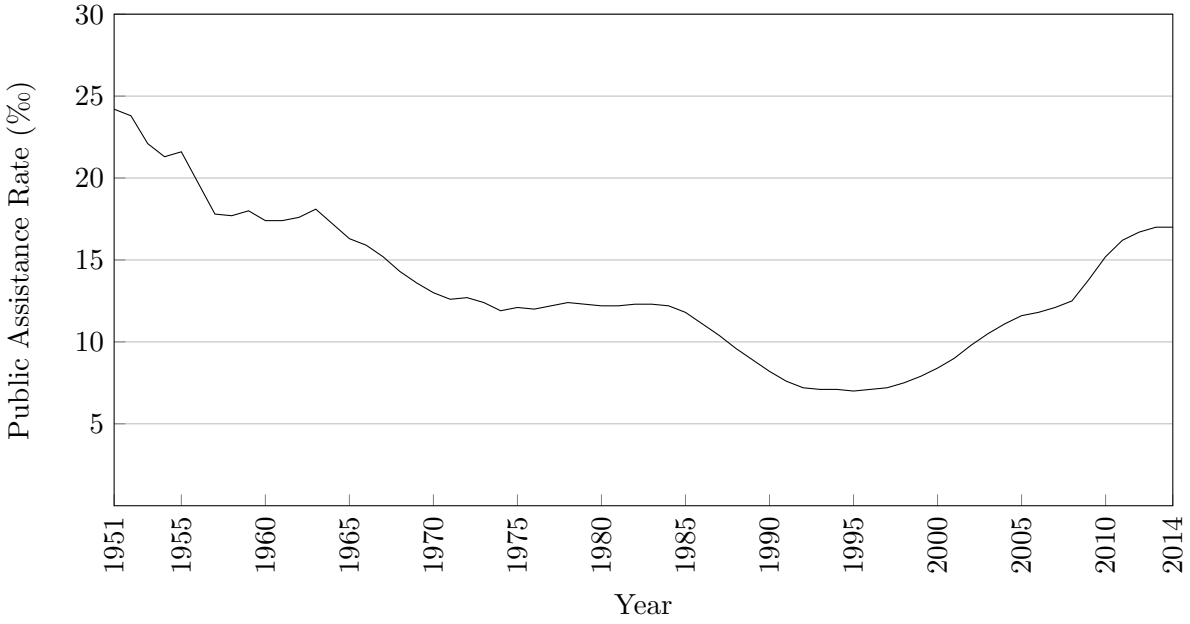

Figure 1: Trend in the Public Assistance Rate in Japan

Sources: Ministry of Internal Affairs and Communications (2012) and Ministry of Health, Labour and Welfare (2010, 2014, 2016b) (Figure made by the author)

Table 5: List of Independent Variables

Independent variables	Description
Age	Age as of 31st January, 2015
Marital status	Dummy: MAR = 1 if married, MAR = 0 otherwise
Sex	Dummy: FEM = 1 if female, FEM = 0 otherwise
Education level (highest completed)	Dummy (base group = less than junior high school): EDU1 = 1 if junior high school; EDU2 = 1 if high school; EDU3 = 1 if junior or specialised training college; EDU4 = 1 if four-year university; EDU5 = 1 if graduate school
Employment status	Dummy: EMP = 1 if employed, EMP = 0 otherwise
Number of children	Children aged 18 or under
Private income	Unit: 10,000 JPY
Other welfare benefits	Unit: 10,000 JPY

Table 6: Descriptive Statistics

Variable	Full Sample ($n = 1,450$)				Participants ($n = 20$)				Non-participants ($n = 1,430$)			
	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max	Mean	Std. Dev.	Min	Max
AGE	56.12	14.38	25	97	61.60	12.85	38	81	56.04	14.39	25	97
MAR	0.80	0.40	0	1	0.25	0.44	0	1	0.81	0.40	0	1
FEM	0.16	0.36	0	1	0.45	0.51	0	1	0.15	0.36	0	1
EDU1	0.10	0.29	0	1	0.40	0.50	0	1	0.09	0.29	0	1
EDU2	0.41	0.49	0	1	0.30	0.47	0	1	0.41	0.49	0	1
EDU3	0.16	0.37	0	1	0.10	0.31	0	1	0.16	0.37	0	1
EDU4	0.31	0.46	0	1	0.15	0.37	0	1	0.31	0.46	0	1
EDU5	0.03	0.18	0	1	0.00	0.00	0	0	0.03	0.18	0	1
EMP	0.70	0.46	0	1	0.30	0.47	0	1	0.71	0.46	0	1
CLD	0.31	0.64	0	6	0.15	0.67	0	3	0.31	0.64	0	6
PRI	671.98	535.31	0	12,389	61.45	68.46	0	179	680.52	534.06	0	12,389
PUB	1.22	12.13	0	179	88.30	55.91	7	179	0.00	0.00	0	0
WEL	6.84	16.41	0	199	1.50	6.71	0	30	6.91	16.49	0	199

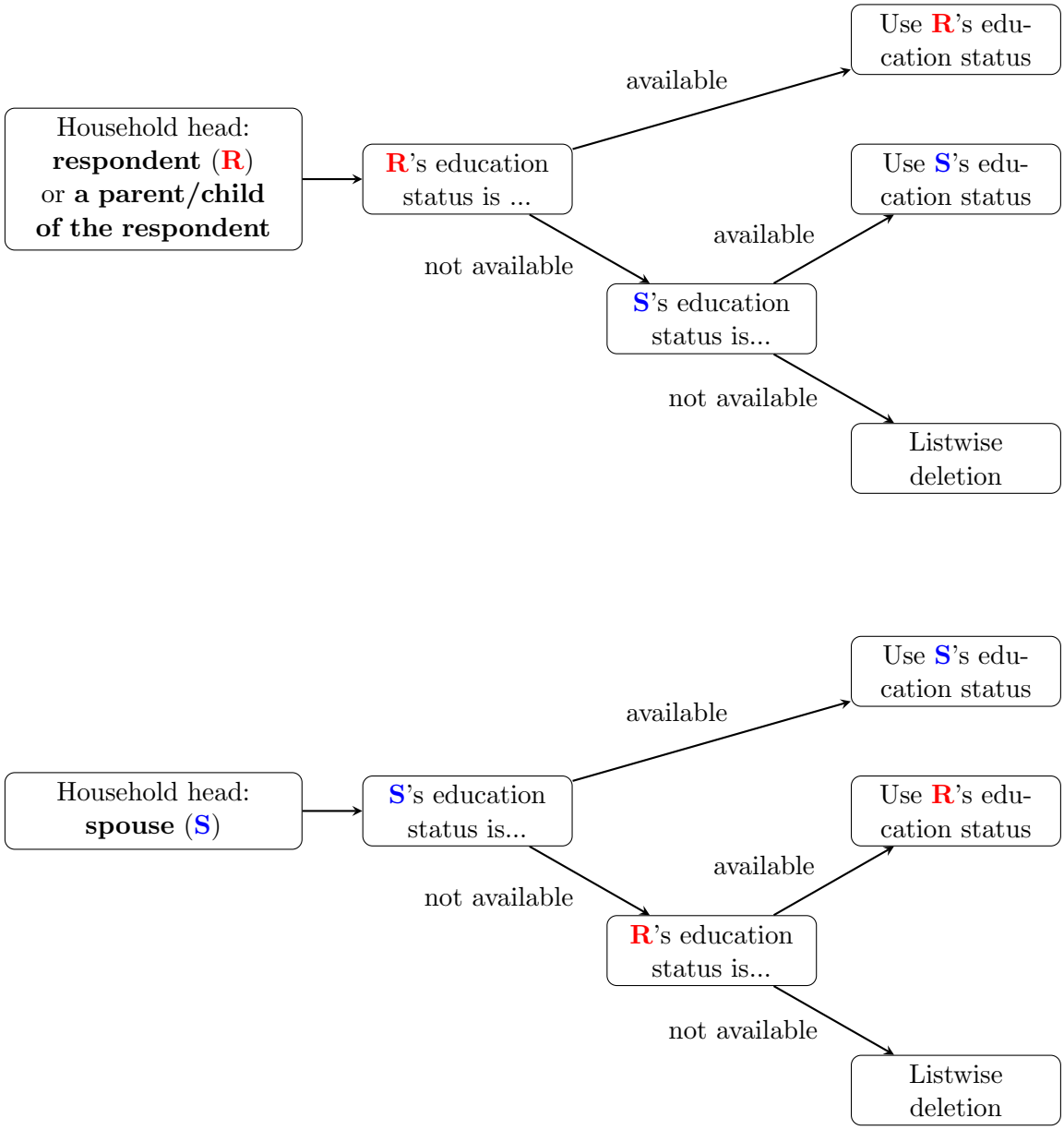

Figure 2: Flow Diagrams for Specification of Education Status

Table 7: Results of Little's MCAR test

Obs	χ^2	df	p-value
1,457	16.033	9	0.066

Table 8: Contiguous Relationship among Regions of Japan

	(1) Hokkaido	(2) Tohoku	(3) Kanto	(4) Chubu	(5) Kinki	(6) Chugoku	(7) Shikoku	(8) Kyushu
(1) Hokkaido		1	0	0	0	0	0	0
(2) Tohoku	1		1	1	0	0	0	0
(3) Kanto	0	1		1	0	0	0	0
(4) Chubu	0	1	1		1	0	0	0
(5) Kinki	0	0	0	1		1	1	0
(6) Chugoku	0	0	0	0	1		1	1
(7) Shikoku	0	0	0	0	1	1		1
(8) Kyushu	0	0	0	0	0	1	1	

Table 9: Bayesian MCMC Estimation Results

	(1) Standard Probit		(2) SAPM		(3) SEPM	
Variable	Coefficient	Std. Dev.	Coefficient	Std. Dev.	Coefficient	Std. Dev.
AGE	0.0033	0.0106	−0.0076	0.0050	0.0002	0.0004
MAR	−0.1182	0.4110	−0.2805	0.2303	−0.0152	0.0159
FEM	0.1810	0.3804	0.2551	0.1610	−0.0066	0.0143
EDU1	0.5270	0.5271	−1.1435	0.8915	−0.0278	0.0358
EDU2	0.0226	0.4501	−1.3456	0.8960	−0.0449	0.0355
EDU3	0.3345	0.5452	−1.2378	0.8767	−0.0488	0.0368
EDU4	−0.0742	0.4926	−0.9328	0.8833	−0.0440	0.0371
EDU5	0.5044	1.1720	−0.7979	0.9200	−0.1901***	0.0596
EMP	−0.1835	0.2882	−0.2143	0.1534	−0.0035	0.0100
CLD	0.2090	0.2929	0.0487	0.0934	0.0072	0.0121
PRI	−0.0119*	0.0006	−0.0004**	0.0002	−0.0002***	0.0000
WEL	−0.0216	0.0162	−0.0078	0.0050	−0.0008	0.0005
α	−0.3265	0.6439	−3.4162***	0.9243	0.0186	0.0465
ρ			−0.0034***	0.0005		
λ					0.2767	0.5120

Note: *** Significant at 1 per cent; ** significant at 5 per cent; * significant at 10 per cent.