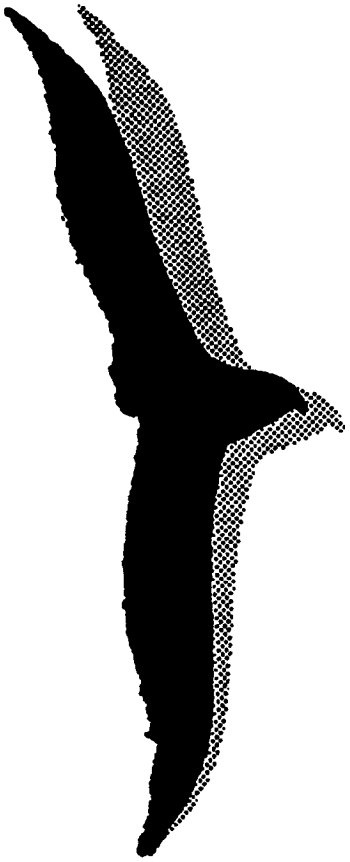
Faculteit der Economische Wetenschappen en **Econometrie**

Serie research memoranda

Estimating the **marginal willingness to pay** for **commuting**

Jos van Ommeren
Gerard J. van den Berg
Cees Corter

Research Memorandum 1998-46 September 1998



/a/l/e/r/t/

applied
labour
economics
research
team

vrije Universiteit *amsterdam*

Estimating the marginal willingness to pay for commuting

Jos van Ommeren *
Gerard J. van den Berg †
Cees Gorter ‡

September 2, 1998

Abstract

With informational frictions on the labor market, hedonic wage regressions provide biased estimates of the willingness to pay for job attributes. We show that a recent theoretical result, which states that variation in job durations does provide good estimates in case of a basic on-the-job search model, can be generalized to a wide class of search models. We apply this result by estimating the marginal willingness of employed workers to pay for commuting, using Dutch longitudinal data. The average willingness to pay for one hour commuting is estimated to equal almost half of the hourly wage rate.

* Cranfield College.
†Department of Economics, Free University Amsterdam, CEPR, and Tinbergen Institute.
‡Department of Economics, Free University Amsterdam, De Boelelaan 1105, NL-1081 HV Amsterdam, The Netherlands, and Tinbergen Institute.

Keywords: job search, commuting distance, job durations, hedonic wages.

1 Introduction

The aim of this paper is to estimate the marginal willingness of employed workers to pay for commuting. There has grown a general interest in the topic of compensation for commutes. Unit commuting costs may be an important determinant of worker behavior. If they are high, then the individual may prefer to reject an offer of a far-away job in favor of a job around the corner even if the former job offers a much higher wage. In that case these costs may affect the allocation process on the labor market. Information on the willingness to pay for commuting distance may help to evaluate policy measures aiming at the abatement of commuting. For example, the direct cost of an additional time unit of commuting due to increased traffic congestion can be calculated.

Usually, the marginal willingness to pay (MWP) for commuting (or for other non-wage job attributes) is estimated by way of hedonic wage regressions (see for example Zax, 1991). This is a natural approach in the context of static or long-run equilibrium models in which markets are assumed to be perfect. In the context of urban economics, the standard monocentric urban residential-location theory implies that wages need to decline with distance from the central business district, to compensate workers located at suburban places for commuting expenses (see, e.g., McDonald, 1997). According to this theory, a labor-market equilibrium locus of wages and commuting times exists, the gradient of the hedonic wage function equals the marginal willingness to pay for commuting time, and these can be estimated directly from the observed relation between commuting time and wages (see Madden, 1995, for an example of an empirical study). The urban theory also implies that workers are (partly) compensated on the housing market for commutes from distant suburbs, because housing prices are lower at higher distances from the central business district.

Gronberg and Reed (1994) and Hwang, Mortensen and Reed (1998) show that estimates obtained from hedonic regressions are likely to be biased if the labor market is characterized by informational frictions. In such markets, firms with a high innate labor productivity offer higher wages as well as better values of the non-wage characteristics, in equilibrium. To the extent that productivity is unobserved, a regression of wages on other characteristics gives bad estimates of the marginal willingness to pay for those characteristics.' In the context of commuting distance, high productivity firms may have a denser net of work locations or they may enable their employees to work at home more often.

'Employing simulated data derived from a rather specific equilibrium search model, Gronberg and Reed (1994) conclude that "the conventional hedonic method generates a MWP estimate that is approximately one-fourth of its true value".

In response to this, Gronberg and Reed (1994) develop a different estimation method for the willingness to pay for job attributes. The starting point in their approach is that workers search on the job in a market with informational frictions. The estimation method exploits the fact that the utility trade-off between the wage and other job attributes is reflected in job duration differences. Specifically, for a basic on-the-job search model, Gronberg and Reed (1994) show that if the job exit rate is much more sensitive to a certain job attribute than to the wage, then this means that the willingness to pay for that attribute is large in absolute value. This approach is not sensitive to the presence of unobserved firm characteristics. Moreover, it is very attractive from a computational point of view, since it suffices to estimate a reduced-form job duration model to estimate the willingness to pay. Note that the latter is a structural parameter as it concerns a characteristic of the instantaneous utility function of the workers.[2]

During the past decades, the search approach has proven to be an useful tool to analyze labor market dynamics in many respects. In empirical work, the focus has been on the determinants of the minimum acceptable wage that induces an individual to accept a job offer (the reservation wage) and the durations of unemployment and jobs. Recently, it has been stressed that "much more attention should also be paid to nonwage characteristics" (Devine and Kiefer, 1993, in their survey on job search). Such non-wage characteristics have mostly been neglected in structural empirical analyses of search models. A noticeable exception concerns Blau (1991) who estimates a job search model with different wage/hours combinations. More recently, Van den Berg and Gorter (1997) analyze a job search model for unemployed individuals that allows jobs to have different wage/commuting-time combinations. In that paper, the structural parameter of interest is the utility trade-off between the wage and commuting time, or in other words the willing-

[2]The general notion that job exits are informative on the workers' willingness to pay for a job attribute has some history. Bartel(l982) studies the effects of the wage and non-wage characteristics on quit decisions, and argues that these effects are informative on the "relative values" (or "relative importance") of the job characteristics for the worker. Herzog and Schlottmann (1990) estimate the effects of the wage and the risk in the workplace on the likelihood that the worker switches to another industry, and they claim that this can be used to assess the willingness to pay for risk reduction. They also argue that the latter may differ from the hedonic (market) price of risk if the labor market is imperfect. However, they do not examine a formal behavioral model for the switching rate. Bartik, Butler and Liu (1992) provide a similar analysis for the housing market, estimating the willingness to pay for neighborhood amenities from residential mobility behavior. See Herzog and Schlottmann (1990) for a listing of other previous literature concerned with the idea that quits are informative on the willingness to pay for attributes. It is interesting to note that the previous empirical studies always find that the estimate of the trade-off between money and other attributes that is based on hedonic regressions is smaller in absolute value than the estimate based on mobility behavior.

ness of unemployed workers to pay for commuting. It is estimated by comparing subjective responses on reservation wages for different job types.

In this paper we apply the approach by Gronberg and Reed (1994) to estimate the marginal willingness to pay for commuting. We use a Dutch dataset that contains information on job durations, job-to-job transitions, and commuting distances of employed workers. In addition to this, in this paper, we generalize the theoretical analysis by Gronberg and Reed (1994) by examining less restrictive dynamic on-the-job search models. We also provide intuition on why the approach cannot be applied in a few particular model extensions.

Section 2 of this paper contains the theoretical analysis. In Section 3 we apply our method to estimate the marginal willingness to pay (MWP) for commuting distance, using information on voluntary job-to-job transitions.[3] We compare the estimates with those from a static hedonic wage regression. Finally, some concluding remarks are made in Section 4.

2 Search theory and the relation between job durations and willingness to pay for job attributes

2.1 The basic on-the-job search model

In this section we examine search models for job-to-job transitions when jobs are characterized by a wage w and a second job attribute x. The formal results do not depend on a specific interpretation of x. However, given the focus of this paper, we will mostly interpret x as the commuting distance, and we will restrict attention to model specifications that make sense under this interpretation of x.

We are particularly interested in the relation between the ratio of the derivatives of the job exit rate θ with respect to w and x on the one hand, and the ratio of the derivatives of the instantaneous utility flow function u with respect to w and x on the other. It has been derived before that in a basic on-the-job search model these ratios are equal to each other (Gronberg and Reed, 1994),

$$\frac{\partial\theta(w,x)/\partial x}{\partial\theta(w,x)/\partial w} = \frac{\partial u(w,x)/\partial x}{\partial u(w,x)/\partial w} \tag{1}$$

[3] An alternative strategy would be to estimate a full structural model for on-the-job search, allowing jobs to have multiple attributes. This would be a formidable task, and, as mentioned above, it is not necessary in order to structurally estimate the MWP.

By definition, the right-hand side of this expression is the marginal willingness to pay for the job attribute x. This is a characteristic of the instantaneous utility flow function u, and as such it is an interesting structural determinant of behavior. The left-hand side of (1) is a quantity that is easily estimated from job duration data. Clearly, therefore, the equality of these ratios enables straightforward estimation of the marginal willingness to pay for x.

We start with a brief description of the basic on-the-job search model. We then generalize this model and examine whether (1) still holds. The theory of on-the-job search aims at describing the behavior of employed individuals who search for a better job (see Mortensen, 1986, and Albrecht, Holmlund and Lang, 1991, for overviews). Consider the basic on-the-job search model extended to allow for non-wage characteristics x, as sketched by Gronberg and Reed (1994). Suppose an individual has a job with characteristics w, x. Offers of new jobs arrive according to a Poisson process with arrival rate λ. Such job offers are random drawings (without recall) from the joint distribution of net wages w^* and job characteristics x^*, with distribution function $F(w^*, x^*)$. We assume that all jobs are full-time jobs. During employment, exogenous separations occur at the rate δ.

Note that we do not assume a parametric functional form for the offer distribution of x^*. In particular, if x^* denotes commuting distance then we allow these to be non-uniformly distributed over space. For example, for an individual living in a village, most job offers may originate from a nearby larger town, so they will have approximately the same commuting distance. We also allow the wage offer to be dependent on the corresponding offer of x^*. For example, in case x^* denotes commuting distance, one may live close to a few small firms offering low wages and far from a town with large firms offering higher wages. Also, firms may provide financial compensation for commuting costs, and the amount of compensation may be increasing in the commuting distance. The latter would establish a positive association between w and x. We do however require that the dependence between w and x is not deterministic, and that they are continuously distributed.[4]

Every time an offer arrives the decision has to be made whether to accept it or to reject it and search further. Individuals aim at maximization of their expected present value of utility over an infinite horizon. We assume that utility is intertemporally separable. The instantaneous utility flow equals $u(w, \text{x})$ in case one works in a job with characteristics w, x.

Individuals are assumed to know λ and $F(w^*, x^*)$. However, they do not know

[4]The continuity assumption is made for notational convenience. If x is discrete (e.g. a dummy variable) then derivatives with respect to x should be replaced by differences.

in advance when job offers arrive, or which w^* and x^* are associated with them. We assume that the model is stationary. This means that w, x, λ and $F(w^*, x^*)$ are assumed to be independent of the duration of being in the present job and independent of all events during the stay in the present job. Further, λ and the function F are not allowed to depend on w or x, and they are assumed to be the same in every job.

For future comparisons, it is useful to present for this model the Bellman equation for the expected present value of utility $R(w, \text{x})$ of someone who works in a job with characteristics w, x. Throughout this section, for convenience, we avoid technicalities and assume regularity conditions to hold. Let ρ be the rate of discount and let U be the expected present value of being unemployed. Using the familiar returns-to-assets representation of Bellman's equation (see e.g. Van den Berg, 1990), we have

$$\rho R(w,x) = u(w,x) + \lambda \int_0^\infty \int_0^\infty \max\{0, R(w^*,x^*) - R(w,x)\} dF(w^*,x^*) + \delta(U - R(w,x)) \tag{2}$$

This equation can be understood by interpreting $R(w, \text{x})$ as an asset for which the return flow equals the flow of what one expects to get from holding the asset. The latter consists of three parts: (i) the instantaneous utility flow, (ii) the job offer arrival rate times the expected gain from finding another job, and (iii) the rate at which a separation arrives times the expected loss of such an event. From equation (2), the individual accepts a job offer if and only if $R(w^*, x^*) - R(w, x) > 0$.

It is well known that in this model, the optimal strategy of employed individuals can be characterized by a reservation utility level $u(w, \text{x})$. A job offer (w^*, x^*) is acceptable if and only if the instantaneous utility flow $u(w^*, x^*)$ associated with it exceeds the reservation utility level $u(w, \text{x})$ (the optimal strategy is "myopic"). Basically, this is because, by accepting an offer, nothing changes except for the increase in the instantaneous utility flow. In other words, no options are thrown away by accepting an offer, and no sunk transition costs are made either. The choice set remains the same when accepting an offer, in the sense that one is always able to return to the previous situation by throwing away part of one's instantaneous utility flow. (Note for future reference that in this model the inequality $R(w^*, x^*) > R(w, \text{x})$ is thus equivalent to the inequality $u(w^*, x^*) > u(w, \text{x})$.)

The exit rate $\theta(w, x)$ out of the present job is the sum of the exit rates to the two different destination states. The exit rate to unemployment equals 6. The exit rate to other jobs equals the product of the job offer arrival rate and the

probability that the offer is acceptable. Let $\mathcal{G}(w, x)$ denote the set of acceptable job offers, i.e.

$$\mathcal{G}(w,x) = \{w^*, x^* | u(w^*, x^*) > u(w,x)\} \tag{3}$$

Note that $\mathcal{G}(w, x)$ depends on w, x solely by way of $u(w, x)$. There holds that

$$\theta(w,x) = \delta + \lambda \int_{\mathcal{G}(w,x)} dF(w^*, x^*) \tag{4}$$

It is clear that $\theta(w, x)$ depends on w and x solely by way of $u(w, x)$. Intuitively, therefore, equation (1) follows. Note that the acceptance probability equals the probability that the random variable $u(w^*, x^*)$ exceeds $u(w, x)$. The probability distribution $F_u(u)$ of $u(w^*, x^*)$ can be obtained from the distribution $F(w^*, x^*)$ of (w^*, x^*). As a result,

$$\theta(w,x) = \delta + \lambda \overline{F}_u(u(w,x)) \tag{5}$$

in which $\overline{F} = 1 - F$. From this, equation (1) immediately follows.[5]

2.2 Generalizations of the basic model

We now examine to what extent the result in equation (1) is robust with respect to some rather unrealistic assumptions of the on-the-job search model.

2.2.1 Endogenous search intensity

Suppose individuals can influence the intensity at which offers arrive (see Mortensen, 1986, and Albrecht, Holmlund and Lang, 1991, for models with endogenous search intensities in case of single job characteristic). Given a particular search effort (or intensity) s, offers of new jobs arrive according to a Poisson process with arrival rate λs. The individual is able to choose s, and if $s > 0$ then he pays a flow of search costs c(s), with c(s) twice differentiable, increasing and convex in s. The instantaneous utility flow equals $u(w - c(s), x)$ in case one works in a job with characteristics w, x and in case the search intensity equals s. Note that we assume search costs to be monetary, i.e. to be paid out of the wage. This assumption will play a crucial role.

[5] A method of proving with wider applicability works by noting that, under regularity conditions, the derivative of θ with respect to w equals the derivative of θ w.r.t. the bounds of the set $\mathcal{G}$ times the derivative of these bounds w.r.t. w. Similarly for x.

In this model version, the set of acceptable job offers is

$$\mathcal{G}(w,x) = \{w^*, x^* | R(w^*,x^*) > R(w,x)\}$$

Note that $\mathcal{G}(w, x)$ depends on w, x by way of $R(w,$ x). This set is not necessarily equal to the set $\mathcal{G}(w,x)$ defined in (3). For example, if w and x in $u(w,$ x) are not perfectly substitutable and if $u(1,\ 0) = u(0,1)$ then a job with characteristics $(w,x) = (1,0)$is. mpe erable over a job with characteristics (w, x) = $(0,1)$ because in the former case s (and therefore the job offer arrival rate Xs) can be higher. In the model, a higher wage income can be allocated towards search activities, whereas a better value of x (a lower commuting distance) cannot. (Note that this assumption could be relaxed somewhat.)

In case of an interior solution for the optimal search intensity $s(w,$ x), the latter follows from differentiation of the equivalent of equation (2) (see the references above). Let u_1 denote the derivative of u with respect to its first argument. We have

$$u_1(w - c(s(w,x)), x) \cdot c'(s(w,x)) = \lambda \int_0^\infty \int_0^\infty \max\{0, R(w^*, x^*) - R(w,x)\}\, dF(w^*, x^*)$$

The exit rate $\theta(w,$ x) now equals

$$\theta(w,x) = \delta + \lambda \cdot s(w,x) \int_{\mathcal{G}(w,x)} dF(w^*, x^*)$$

Clearly, in general, $s($ $w,$ x) does not depend on w, x only by way of $u(w,$ x) , and, therefore, neither does $\theta(w,x)$. This is due to the fact that in general u_1 (w – c(s) $,x$) does not depend on w, x by way of $u(w,x)$ (take e.g. a Cobb-Douglas specification for u). Note that Gronberg and Reed (1994) erroneously state that (1) does hold if search intensity is endogenous.

Now let us make the simplifying assumption that preferences are additive and linear in w and x (so $u(w,$ x) can be written as w + ax). This means that w and x are perfect substitutes; x is like money. From the perspective of a job searcher, search costs can now be paid as effectively out of w as out of x. It is not difficult to see that, as a result, the set of acceptable job offers is as in (3). Moreover, u_1 is a constant. Because of this, both $s(w,x)$ and $\theta(w,x)$ depend on w, x solely by way of $u(w,$ x), and the model can be rewritten as a model with a single job characteristic u. As a result, (1) does hold. Note that Gronberg and Reed (1994) take a nonlinear specification for $u(w,$ x) in their empirical model, so their results are only valid under the interpretation that search intensity is exogenously determined.

2.2.2 Business cycles

Suppose that the business cycle affects the current values of the structural determinants λ and F but not the current values of w or x of a given worker. (Burgess, 1989, presents such a model in case of a single job characteristic.) This model is nonstationary in the sense that the expected present utility value $R(w, x)$ now varies over calendar time. However, it is not difficult to see that the optimal strategy is still myopic (and stationary): a job offer (w^*, x^*) is acceptable if and only if the instantaneous utility flow $u(w^*, x^*)$ associated with it exceeds the reservation utility level. By accepting an offer, nothing changes except for the increase in the instantaneous utility flow. In particular, the search environment in the new job and the way it changes over the business cycle are exactly the same as in the old job. As a result, equation (1) is still valid.

Note that it does not matter whether the individual anticipates the business cycle effects or not, as long as this does not change over time. Also note that the model in which F varies over calendar time is formally equivalent to a model in which general human capital is accumulated but the rents of it can only be extracted by the worker in case a new job is accepted.

In the model considered here, λ and/or F vary over time. As a result, the job exit rate is duration dependent, which means that $\theta(w, x)$ depends on the elapsed job duration t. However, obviously, the ratios in the equality (1) do not depend on t.

2.2.3 Limited maximum number of transitions

The basic on-the-job search model does not impose an upper bound on the number of job-to-job transitions by a worker in any given time interval. Consider the opposite situation in which an employed worker can change jobs at most once. This model resembles the standard job search model for unemployed workers (Mortensen, 1986), since in the latter model it is assumed that a job is kept forever, so an unemployed individual can only make one transition.

In our case the employed worker with a job (w, x) accepts a job offer (w^*, x^*) if and only if the instantaneous utility $u(w^*, x^*)$ of the offered job exceeds $(\rho + \delta)R(w, \text{x}) - \delta U$ (this follows from a comparison of the present value of accepting the job offer and the present value $R(w, \text{x})$ of continuing search). The latter quantity is the reservation utility level $\xi(u(w, x))$ associated with a job (w, x). Let $F_u(u)$ denote the probability distribution of $u(w^*, x^*)$. The job exit rate can then be expressed as $\delta + \lambda\overline{F}_u(\xi(u(w, \text{x})))$ (note the analogy to equation (5)). As a result, equation (1) is still valid.

2.2.4 Transaction costs

Suppose that, every time one moves from one job to another, an amount of money c has to be paid instantaneously. Hey and McKenna (1979) and Van den Berg (1992) analyze on-the-job search models with such "transaction costs", under the assumptions that jobs are fully characterized by wages and individuals are risk-neutral. In order to maintain stationarity we assume that c does not depend on the time spent in the present job nor on events during the stay in the present job.

In this model it is not optimal to accept every job offer with a higher instantaneous utility flow. Because of the costs to be paid at every transition, there is an incentive to reduce the number of transitions, i.e. to be more selective with respect to job offers (see the references above).

In case of risk-aversion, it is difficult to analyze such a model. The transaction costs have to be paid out of savings, so it is necessary to include an asset equation to the model. However, intuitively it is clear that the equality (1) fails to hold here in general. Basically, a higher wage is more attractive than a better value of x, because the transaction costs have to be paid out of future wages.

If preferences are additive and linear in w and x then payment of c can be thought of as coming out of a sufficiently large amount of given wealth. In addition, analogously to the model in Subsection 2.2.1 with additive and linear preferences, x is now like money, and transaction costs can be paid as effectively out of w as out of x. Indeed, $\theta(w, x)$ now depends on w, x solely by way of $u(w$, x), and the model can be rewritten as a model with a single job characteristic u. As a result, (1) does hold.

2.2.5 Other extensions

Together, the sub-subsections above provide the following insight. If (i) the instantaneous utility function is not additive and linear in w and x, while (ii) the search environment is not symmetric in w and x in the sense that having a high w is more useful for further search than having a good value of x, then the equality (1) fails. Basically, in such cases, it is more beneficial to have a high wage, since part of that wage can be dedicated to improved search relatively easily.

This insight is confirmed by examining other model extensions. Sometimes however the asymmetry in search environment is so strong that additivity of the utility function does not help. Consider a model in which, in addition to job offers, there are also offers of just x, arriving according to a second Poisson process. Such a model has some relevance in case x denotes the commuting distance, since a worker in a given job with a given w may from time to time have the opportunity

to move his residence towards a location that is closer to his work (in reality, he may of course also move for other reasons). In this model, the search environment is not symmetric in w and x: it is easier to search for x than for w. In such a case a job with a high wage and a high commuting distance is more attractive than a job with a low wage and low commuting distance, because it is relatively easy to improve after a while on a high commuting distance. A worker may thus even accept a job offer with a somewhat lower instantaneous utility flow, if such a job is characterized by a high wage and a high commuting distance. Note that the value of x need not be constant within a job spell in this model. This further complicates attempts to proceed along the lines of Subsection 2.1 here.[6]

The model above suggests that the method of Subsection 2.1 for making inferences on the willingness to pay for commuting distance is sensitive to the assumption that workers do not move their residence somewhere within a job spell. This method may therefore be best suited to study job mobility of workers for which the latter assumption is likely to hold, such as members of settled households, as opposed to single-living schoolleavers.

2.3 A general model framework

In this subsection we provide a general on-the-job search model framework that incorporates some of the generalizations of Subsection 2.2 as well as some other generalizations. We show that equation (1) holds in general, and we consider the effects of the current values of w and x on the expected present value.

Consider the model of Subsection 2.1, with the modification that here we do not make any assumption on the search environment after the first job-to-job transition that will be made. Denote the expected present value of moving to a job with characteristics w^*, x^* by $\tilde{R}(w^*, x^*)$. We do not restrict the shape of $\tilde{R}$ as a function of w^*, x^*. A job offer is acceptable if and only if $\tilde{R}(w^*, x^*) > R(w, x)$.

Let $\mathcal{G}(w, x)$ denote the set of acceptable job offers, i.e.

$$\mathcal{G}(w, x) = \{w^*, x^* | \tilde{R}(w^*, x^*) > R(w, x)\}$$

Note that $\mathcal{G}(w, x)$ depends on w, x by way of $R(w, x)$. The job exit rate $\theta(w, x)$ is still described by

$$\theta(w, x) = \delta + \lambda \int_{\mathcal{G}(w,x)} dF(w^*, x^*)$$

[6] Other examples for which (1) fails to hold concern models in which the current value of x affects the wage offer distribution $F(w^*, x^*)$ or the job offer arrival rate λ.

although the definition of $\mathcal{G}$ is different from the definition in equation (4). Furthermore, the expected present value $R(w, x)$ of being in a job with characteristics w , x satisfies

$$\rho R(w, x) = u(w, x) + \lambda \int_{\mathcal{G}(w,x)} \tilde{R}(w^*, x^*) - R(w, \mathrm{x})\, \mathrm{d}F(w^*, x^*) + \delta(U - R(w, x)) \quad (6)$$

It is useful to examine which of the models of the previous subsections are special cases of the present model. First of all, the basic model of Subsection 2.1 is a special case of the current model (take $\tilde{R} \equiv$ R). Secondly, the models with endogenous search intensities can not be rewritten as the current model. The same is true for the nonstationary model with business cycle effects. However, it is straightforward to include general nonstationarity in the model of this subsection. In that case, the derivatives of $\theta(w, x)$ with respect to w or x may depend on the elapsed job duration. The model with one possible job-to-job transition can be rewritten in terms of the model of this subsection, by taking $\tilde{R}(w^*, x^*) = (u(w^*, x^*) + \delta U)/(\rho + \delta)$. The model with transaction costs and an additive linear utility function can also be rewritten as the current model, by taking $\tilde{R}(w^*, x^*) = R(w^*, x^*) -$ c. Models in which the job offer arrival rate or the wage offer distribution are different during the next job are also special cases of the current model.

The function $\theta(w, \mathrm{x})$ depends on w and x solely by way of $R(w, x)$. Therefore, under regularity conditions, by analogy with the derivations in Subsection 2.1 it follows that

$$\frac{\partial \theta(w, x)/\partial x}{\partial \theta(w, x)/\partial w} = \frac{\partial R(w, x)/\partial x}{\partial R(w, x)/\partial w} \quad (7)$$

Moreover, by differentiation of equation (6) with respect to w and with respect to x it follows that

$$\frac{\partial R(w, x)/\partial x}{\partial R(w, x)/\partial w} = \frac{\partial u(w, x)/\partial x}{\partial u(w, x)/\partial w} \quad (8)$$

By combining equations (7) and (8) it follows that equation (1) holds under quite general conditions. The left-hand side of (8) compares the relative importance of the current values of w and x for the expected lifetime utility. This term could be called the lifetime MWP, whereas the right-hand side of (8) could be called the *instantaneous MWP*. Interestingly, these two MWP measures have the same value, so the distinction between "instantaneous" and "lifetime" is irrelevant.

We conclude Section 2 by summarizing the main results. First, the equality (1) holds in more general settings than just the basic on-the-job search model.

Secondly, there are also model extensions for which (1) fails in general. However, if utility is additive and linear in the two job attributes then (1) does hold in a number of these extensions. So, it appears that in this respect there is a trade-off between assumptions on the search environment and assumptions on the utility function. Thirdly, in general, the relative importance of the two attributes for lifetime utility is equal to the relative importance in the instantaneous utility flow, and therefore the ratio of derivatives of the job exit rate with respect to the two attributes also captures the relative importance of these attributes for lifetime utility.

3 The empirical analysis

3.1 The data

In the empirical analysis we use data from the so-called Telepanel dataset. This is a survey held among households in The Netherlands. For our purposes, the main advantage of this dataset is that it contains information on the location of the workplace and the residence, and thus, by implication, on the commuting distance. Since the dataset has been described elsewhere numerous times and in great detail (see Van Ommeren, 1996, for an overview), the current exposition will be very brief. We use data that were collected in 1992-1993. These contain extensive retrospective information on the life course history of the respondents, notably concerning the histories of labor market behavior and the behavior concerning the residence, and the changes in household composition. In particular, the data record the starting and ending dates of the spells spent in different labor market states (notably unemployment and employment; job-to-job transitions within a spell of employment are recorded as well). The data thus enable observation of the durations spent in these states (including job durations). In addition, for each job, a number of job characteristics are recorded. The data on job exits allow for a distinction between voluntary quits and involuntary lay-offs.

We select the male respondents who (i) worked for more than 20 hours per week on January 1, for at least one of the years 1985-1991, and (ii) for which all relevant variables (job duration, wage, commuting distance, and other background characteristics) are observed. This results in a sample of 370 individuals. From this sample, 318 individuals work for more than 20 hours per week at the first annual inspection point (January 1985). The other 52 individuals flow into such a job at a later moment. Among the 318 individuals, the mean elapsed job duration in January 1985 equals 4.6 years. In total, we observe 636 job spells in our sample. Of these, 270 end in a transition to another job. The remaining are

either right-censored or end in a transition out of employment.

We observe the municipalities of the residence and the workplace of the worker (more detailed information on commuting distance or commuting time is absent), and we use distance between the center of the municipalities as our commuting distance variable. This variable under-estimates the real distance in case one lives and works at the same municipality. It is likely to over-estimate the distance in the other case, as workers can be expected to self-select such that they live and work in those parts of the municipalities that face each other. The empirical reduced-form job-duration model that we estimate below includes a range of other explanatory variables. These represent job-, worker-, and labor market characteristics. Specifically, we include age (in classes), size of branch, number of subordinates, civil servant, on payroll, full-time employed, sector (construction), and educational level (university, polytechnic, vocational, high school, lower vocational). The time-varying variables are allowed to change yearly. Biannual dummy variables are incorporated to capture changes in general labor market conditions. Table 1 lists the sample means of all explanatory variables included in the empirical model.

3.2 The empirical model specification and the likelihood function

Recall that we aim to estimate $(\partial\theta/\partial x)/(\partial\theta/\partial w)$, i.e. the ratio of the marginal effects of commuting distance x and the wage w on the job exit rate θ. We assume that the way in which the job-to-job transition rate depends on its determinants can be captured by a Mixed Proportional Hazard model (Lancaster, 1990). In this model we assume a constant baseline hazard. Hence, we exclude autonomous duration dependence, in line with most of the theoretical models of Section 2. For the individual job exit rate θ this means that we can write (in notation explained below),

$$\theta = \delta + \exp(\beta' z) \, . \, v$$

where z is a vector with the explanatory variables, β is a vector of unknown parameters, and v is an unobserved variable capturing unobserved worker-specific heterogeneity. Obviously, the wage w and commuting distance x are included in z. We assume that v is independent of z. Inclusion in the model of unobserved heterogeneity is necessary to avoid inconsistent estimation of β (see e.g. Lancaster, 1990).

In line with the literature, we use the log wage as a regressor in z. In fact, we investigated the functional forms of the effects of w and x on the exit rate to other

jobs by estimating separate models with linear and log-linear transformations. It turns out that for the wage, a log-linear transformation fits the data better than a linear transformation. With respect to commuting distance, there is no significant difference between the two specifications. The results below are based on a linear specification.

Let β_i denote the element of β corresponding to variable i. In particular, β_{logw} denotes the coefficient of logw. The MWP for x equals

$$\text{MWP}(x) = \frac{\beta_x}{\beta_{logw}} \cdot \text{w} \tag{9}$$

Recall that the data are informative on the destination state upon job exit. Under the maintained assumption that v does not affect δ, the estimation of the job-to-job transition rate can be performed separately from the estimation of δ. Spells ending in a transition to unemployment are then treated as right-censored observations of the duration until a job-to-job transition (Lancaster, 1990). One may be tempted to think that the parameter δ is allowed to depend on z, by analogy to the standard result in duration analysis that the moment of right-censoring is allowed to depend on observable conditioning variables. However, if δ depends on w or x then the search-theoretical predictions from Section 2 on the effects of w and x on the job-to-job transition rate may be violated. Note that this remark is of importance for any empirical analysis of job-to-job transitions.

We assume that the distribution of v is discrete. Initially we assume that v has two mass points. We denote these mass points by v_1 and v_2, and their probabilities in the inflow into jobs by p_1 and p_2, respectively. This distribution is flexible and attractive from a computational point of view.

We estimate the model by Maximum Likelihood. Consider the available information on job durations. As noted in the previous subsection, some individuals have an ongoing job spell on January 1, 1985, whereas others flow into a job afterwards. In both cases, we may observe multiple job spells for a given individual. We include such multiple job spells in the analysis for the reason that they increase the sample size of job spells, and in particular because, as is well known, the model estimates are less sensitive to the proportionality assumption of the job-to-job transition rate than if only single spells are used (see Honor-e, 1993). Basically, this exploits the assumption that the unobserved heterogeneity term v is constant across jobs. However, it should be noted in advance that estimation without the use of these spells does not affect the results in any important way.

Now consider the derivation of the individual likelihood contributions. We start with the individuals who have an ongoing job spell on January 1, 1985.

Derivation of the distribution of the length of this spell (conditional on the characteristics of the job) is not straightforward. For example, in a very general setting, this distribution may depend on the distribution of the job characteristics (including w and x) across jobs, and on the way in which the distribution of acceptable values of such characteristics changes over consecutive jobs.[7] To keep the analysis manageable, we proceed in a rather ad-hoc way. First of all, we condition on the elapsed job duration p at the moment of sampling. If instead we would have used the unconditional job duration t as an endogenous variable, then we would have to make a number of strong untestable assumptions on the inflow rate into jobs before the sampling moment, and the results would depend on this (see e.g. Heckman and Singer, 1984, and Ridder, 1984). The population distribution of the job duration until exit into another job given z and v is exponential, with parameter $\exp(\beta' z)v$. Let $g(t|z,v)$ and $\overline{G}(t|z,v)$ denote the corresponding population density and survival function, respectively. Under some assumptions, the density h of $t|p,\ z$ equals

$$h(t|p,z) = \frac{\mathrm{E}_v\ [g(t|z,v)]}{\mathrm{E}_v\ [\overline{G}(p|z,v)]} \quad \text{with } t \geq p$$

where the expectations are taken with respect to the distribution of v in the population. The expression above takes account of the "length-biased sampling" phenomenon (see e.g. Ridder, 1984): the distribution of durations of ongoing spells dominates the population duration distribution. Moreover, individuals with small v have on average longer t.

If this ongoing spell is the only spell for this individual in the data then the likelihood contribution is equal to $h(t|z,)$ above. Modification in case of right-censoring is straightforward. Now suppose that we observe n + 1 job spells for an individual. It is not difficult to show that the likelihood contribution then equals

$$\frac{\mathrm{E}_v\ \left[\ g(t_1|z_1,v)\cdot\ \prod_{j=2}^{n+1}\ g(t_j|z_j,v)\ \right]}{\mathrm{E}_v\ \left[\ \overline{G}(p|z_1,v)\ \right]} \quad \text{with } t \geq p$$

where the index oft and z denotes the chronological-ordering number of the spell. Again, modification in case of right-censoring of a job duration is straightforward.

Now consider an individual whose first job spell starts after January 1985. Under weak conditions, the duration distribution for such a job spell is equal

[7] This can be seen by interpreting the job spells as outcomes of a stationary stochastic decision process with unobserved heterogeneity, incorporating the workers' optimal strategy. For example, for a given individual, the values of v and w at a given date may be related because individuals with high v move on quickly to jobs with a high w.

to the population distribution of job durations. If we observe multiple job spells for this individual then the likelihood contribution simply equals the expectation over v of the product of $g(t_j|z_j, v)$.

During the period of observation, the values of the explanatory variables z may change for a given individual. To the extent that such changes correspond to job changes, the model specification and the likelihood above take account of them. However, the values of z can also change within a job spell. One may question whether such changes should be included in the empirical model. On the one hand, they may constitute a potentially important empirical determinant of job exit behavior. On the other hand, allowing z to vary over time would mean that the empirical model is not in full agreement to the basic theoretical model anymore (recall that the job exit rate in the latter model is constant over time). However, the theoretical model extension discussed in Subsection 2.2.2 allows for time-varying determinants of the job exit rate, and the main predictions of the basic model remain valid under this extension. Alternatively, one may think of changes in z as unanticipated shocks in the search environment, in which case the main predictions remain valid as well. We therefore decide to take a compromise stand by allowing the explanatory variables (like age group) only to change annually. In particular, we adopt the values at January 1 for the whole calendar year.

3.3 Estimation results

The estimates of the parameters of the job-to-job transition rate are in Table 1. The most important estimates are those for the wage and commuting distance effects, since these serve as inputs for the estimate of the marginal willingness to pay for commuting. Table 1 shows significant effects for both: higher wages reduce the job-to-job transition rate whereas higher commuting distances increase this rate. More specifically, the estimated quasi-elasticity of the expected duration until exit to another job with respect to commuting distance per kilometer $(\beta_x/100)$ is equal to -0.0038, and the estimated elasticity with respect to the wage (β_{logw}) is equal to 1.49. The estimate of the MWP for commuting follows from the right-hand side of equation (9). We substitute the estimated β_x and β_{logw}, we substitute for w the average net hourly wage of 20 Dutch Guilders, and we divide the estimate by 100 in order to obtain the willingness to pay for one kilometer. The resulting estimate equals -0.051, with a standard deviation of 0.028.[8] As a consequence, the MWP in case of a working day of 8 hours is es-

[8] If we allow for a third mass point in the heterogeneity distribution then the estimate increases slightly (to 0.058). Exclusion of unobserved heterogeneity v from the model leads to

timated to be about -0.40 Guilder per kilometer. Since the average commuting distance is 20 kilometers, the over-all average MWP in case of a working day of 8 hours is estimated to be 8 Guilders.

The estimated MWP for commuting distance can be used to estimate an MWP for commuting *time*. The annual Dutch labor force survey, called "Enquête Beroepsbevolking" contains information that can be used to calculate the average traveling speed during commuting (the speed of course depends on the mode used). According to Statistics Netherlands (1992), commuters who use the car travel with an average speed of about 32 kilometers per hour for commutes of less than 16 kilometers. Commuters who travel for more than 16 kilometers travel on average twice as fast. Based on these figures, the MWP for one-hour commuting per day (i.e., two half-hour car trips per day) is about minus one-third of the hourly wage rate,[9] while the MWP for more than one hour commuting per day (i.e., two trips per day that each take more than half an hour) is about minus two-third of the hourly wage rate. Since the average commuting distance is 20 kilometers, the over-all average is closer to minus one-third of the wage rate than to minus two-third of the wage rate. These estimates are well in line with other empirical results (see Wales, 1978, who, in the context of a labor supply model, estimates that commuting time is on average valued at about two-third of the hourly wage rate, and Small, 1992, who concludes from a large number of studies that the average value of time for the journey-to-work trip is estimated to be around 50 percent of the hourly wage rate).

Most other covariate effects are in line with those found in previous empirical studies on job and residential mobility (see, for example, Lindeboom and Theeuwes, 1991, Van den Berg, 1992, and Van Ommeren, 1996). The calendar year effects reflect aggregate fluctuations in job mobility in the Dutch labor market. We observe a general recovery of labor market conditions in the second half of the eighties. Not surprisingly, older workers are less mobile than middle-aged (34-44 years) workers. Workers in large firms are less mobile, perhaps because there are more internal career possibilities within large firms.

It may be interesting to examine to what extent the MWP estimate is biased if the model of Subsection 2.2.5 holds, i.e. if workers sometimes receive the opportunity to move their residence during a job spell. We know that in that model, equation (1) does not necessarily hold, in which case the MWP is not equal to

an even higher estimate (0.068), but this model gives a significantly worse fit than the model with unobserved heterogeneity.

[9] At a speed of 32 km/hour, a commuting time of half an hour corresponds to a commuting distance of 16 kilometers. The latter costs 16 times 0.40 Guilder (i.e. 6.4 Guilders) per day. This is about one third of the average hourly wage of 20 Guilders.

the right-hand side of (9). Indeed, the value of x associated with a given job offer is of less importance to the worker. As a result, the job-to-job transition rate will be less sensitive to the current value of x, and the instantaneous MWP will be under-estimated.

Finally, we compare our estimates with those from a conventional hedonic wage regression. As noted in the introduction, this approach, which has been popular in the literature, is likely to produce biased estimates of the MWP. We simply regress the log hourly wage rate on the same set of regressors as used above, so $\log w = \alpha' z + \varepsilon$. The estimation results are in Table 2. Note that the firm size has a positive effect on the wage. This is in line with the equilibrium search models used as a motivating theoretical framework by Gronberg and Reed (1994) and Hwang, Mortensen and Reed (1998). Our focus is on the estimate for commuting distance, which is significantly positive and equal to 0.0010 per kilometer. According to the standard compensating-wage model with perfect markets, the MWP for commuting equals $-\partial w/\partial x$. Evaluated at the average hourly wage of 20,[10] the estimated MWP is equal to about -0.02, which is less than half of the MWP estimate based on the job duration data.

It is important to note that the ratio of the MWP estimates from the duration model and the hedonic model (this ratio equals 2.5) is robust against the functional form of the way in which the wage rate and the commuting distance enter these models. We conclude from this that the hedonic model underestimates the MWP for commuting.

4 Conclusion

We have shown that a recent theoretical result, which states that variation in job durations can be exploited to estimate the willingness to pay for job attributes, can be generalized. The theoretical result was derived in the context of a basic on-the-job search model, and in the paper we show that a number of assumptions of the basic model can be relaxed, and, indeed, that the result is valid in a wide class of search models. In addition, we show that the relative importance of the current job attributes for the expected lifetime utility is equal to the relative importance for the instantaneous utility flow. So, for the marginal willingness to pay, the distinction between "instantaneous" and "lifetime" is irrelevant.

In the application, we estimate a job duration model using Dutch longitud-

[10] The estimated regression coefficient equals the quasi-elasticity $(\partial \log w/\partial x)$, and hence needs to be multiplied with the average hourly wage and divided by 100 to obtain minus the average MWP.

inal data. The estimation results are subsequently used to estimate the marginal willingness of employed workers to pay for commuting. The MWP in case of a working day of 8 hours is estimated to be about -0.40 Guilder per kilometer (or $ – 0.20 per kilometer, or $ – 0.32 per mile), per day. Since the average commuting distance is 20 kilometers, the over-all average MWP in case of a working day of 8 hours is estimated to be 8 Guilders per day. These estimates can be translated into an MWP for commuting time. It turns out that the over-all population average of the marginal willingness to pay for a reduction of commuting time is almost half the hourly wage rate. This suggests that commuting is not experienced as a complete waste of time.

References

Albrecht, J.W., B. Holmlund, and H. Lang (1991), "Comparative statics in dynamic programming models with an application to job search", *Journal of Economic Dynamics and Control, 15, 755-769.*

Bartel, A.P. (1982), "Wages, nonwage job characteristics, and labor mobility", *Industrial and Labor Relations Review, 35, 578-590.*

Bartik, T.J., J.S. Butler and J.T. Liu (1992), "Maximum score estimates of the determinants of residential mobility: implications for the value of residential attachment and neighborhood amenities", *Journal of Urban Economics, 32,* 233-256.

Blau, D.M. (1991), "Search for nonwage job characteristics: a test of the reservation wage hypothesis", *Journal of Labor Economics, 9,* 186-205.

Burgess, S.M. (1989), "The estimation of structural models of unemployment duration with on-the-job search', Working paper, University of Bristol, Bristol.

Devine, T.J. and N.M. Kiefer (1993), "The empirical status of job search theory", *Labour Economics,* 1, 3-24.

Gronberg, T. J. and W.R. Reed (1994), "Estimating workers' marginal willingness to pay for job attributes using duration data", *Journal of Human Resources, 29, 911-931.*

Heckman, J. J. and B. Singer (1984), "Econometric duration analysis", *Journal of Econometrics, 24, 63-132.*

Herzog, H.W. Jr. and A.M. Schlottmann (1990), "Valuing risk in the workplace: market price, willingness to pay, and the optimal provision of safety", *Review of Economics and Statistics, 72, 463-470.*

Hey, J.D. and C.J. McKenna (1979), "To move or not to move", *Economica, 46,* 175-185.

Honoré, B.E. (1993), "Identification results for duration models with multiple spells", *Review of Economic Studies, 60, 241-246.*

Hwang, H.S., D.T. Mortensen and W.R. Reed (1998), "Hedonic wages and labor market search", *Journal of Labor Economics,* forthcoming.

Lancaster, T. (1990), *The Econometric Analysis of Transition Data,* Cambridge University Press, Cambridge.

Lindeboom, M. and J. Theeuwes (1991), "Job duration in The Netherlands", *Oxford* Bulletin *of Economics and Statistics, 53, 243-264.*

Madden, J.F. (1985), "Urban wage gradients: empirical evidence", *Journal of Urban Economics,* 18, 291-301.

McDonald, J.F. (1997), *Fundamentals of Urban Economics,* Prentice Hall, New Jersey.

Mortensen, D.T. (1986), "Job search and labor market analysis", in 0. Ashenfelter and R. Layard, editors, *Handbook of Labor Economics,* North-Holland, Amsterdam.

Ridder, G. (1984), "The distribution of single-spell duration data", in G.R. Neumann and N. Westergard-Nielsen, editors, *Studies in Labor Market Analysis,* Springer-Verlag, Heidelberg.

Small, K.A. (1992), *Urban Transportation Economics,* Harwood, Chur.

Statistics Netherlands (1992), *Enquête Beroepsbevolking* 1992, Statistics Netherlands, Heerlen.

Van den Berg, G.J. (1990), "Nonstationarity in job search theory", Review of *Economic Studies, 57, 255-277.*

Van den Berg, G.J. (1992), "A structural dynamic analysis of job turnover and the costs associated with moving to another job", *Economic Journal,* 102, 1116-1133.

Van den Berg, G.J. and C. Gorter (1997), "Job search and commuting time", *Journal of Business and Economic Statistics,* 15, 269-281.

Van Ommeren, J.N. (1996), "Commuting and relocation of jobs and residences; a search perspective", Ph.D. thesis, Thesis Publishers, Amsterdam.

Wales, T.J. (1978), "Labour supply and commuting time; an empirical study", *Journal of Econometrics, 8,* 215-226.

Zax, J.S. (1991), "Compensation for commutes in labor and housing markets", *Journal of Urban Economics, 30,* 192-207.

Table 1 **ESTIMATES of coefficients for voluntary job-to-job mobility [a]**

variables[b]	β duration model			mean value
age:				
< age < 24	0.27	(0.31)		0.08
24 < age < 34	0.44	(0.26)	*	0.38
34 < age < 44	0.79	(0.27)	**	0.31
size of branch:				
size > 200 p	-0.57	(0.20)	**	0.39
20 p < size < 200 p	-0.32	(0.20)		0.33
number of subordinates:				
0	-0.20	(0.28)		0.19
1,2,3	0.13	(0.19)		0.68
non government	0.28	(0.22)		0.65
on payroll	0.19	(0.25)		0.97
construction sector	-0.16	(0.35)		0.07
more than 32 hours	-0.44	(0.28)		0.97
ln(wage rate)[c]	-1.49	(0.23)	**	2.86
educational level:				
university	0.33	(0.29)		0.10
polytechnic	0.21	(0.26)		0.16
vocational	-0.45	(0.31)		0.23
lower vocational	-0.75	(0.30)	**	0.26
high school	-0.05	(0.64)		0.07
commuting distance[d]	0.38	(0.20)	*	0.20
calendar year:				
1985/1986	-0.36	(0.18)	**	0.28
198711988	-0.10	(0.17)		0.29
1989/1990	0.00	(0.08)		0.29
unobserved heterogeneity: mass points and probabilities":				
v_1	4.07	(4.19)		
v_2	13.84	(4.26)	**	
P_1	0.99	(0.01)		
P_2	0.01	(0.01)		

Number of observations	370
Log-likelihood	-701.85

(a) Standard errors in parentheses; **: significant at 5%, *: significant at 10%.

(b) Reference groups: age (older than 44), size of branch (less than 20), number of subordinates (more than 3), non government (civil servant), on payroll (selfemployed), more than 32 hours (less than 32), construction sector ("others"), educational level (primary and "others"), calendar year (1991).

(c) Net wage (in Dutch guilders) per hour; 2 guilders is about 1 US $.

(d) Distance in 100 kilometres.

(e) Inclusion of more masspoints hardly changes the estimates.

Table 2 **ESTIMATES of hedonic wage regression [a]**

variable&	a wage regression[b]		
constant	3.33	(0.07)	*
age:			
< age < 24	-0.55	(0.03)	*
24 < age < 34	-0.27	(0.02)	*
34 < age < 44	0.07	(0.02)	*
size of branch:			
size > 200 p	0.23	(0.02)	*
20 p < size < 200 p	0.10	(0.02)	•
number of subordinates:			
0	0.08	(0.02)	*
1,2,3	0.02	(0.02)	
non-government	0.02	(0.02)	
on payroll	-0.13	(0.04)	*
more than 32 hours	-0.36	(0.04)	*
construction sector	-0.14	(0.03)	*
educational level:			
university	0.24	(0.03)	*
polytechnic	0.14	(0.02)	*
vocational	0.01	(0.02)	
lower vocational	-0.07	(0.02)	*
high school	0.02	(0.03)	
commuting distance[d]	0.10	(0.03)	*
calendar year:			
1985/1986	0.01	(0.02)	
1987/1988	-0.01	(0.02)	
1989/1990	-0.02	(0.02)	
Adjusted R^2 = 0.32			
Number of observations = 2332"			

(a) Standard errors in parentheses; *: significant at 5%.
(b) Dependent variable is the logarithm of the net wage per hour (in Dutch guilders).
(c) Reference groups: as in Table 1.
(d) Distance in 100 kilometres.
(e) Each individual is included once for every year, provided he works on the first of January in that year.