Identification of 'Wasteful Commuting' using Search Theory

Jos Van Ommeren[1]
Willemijn Van der Straaten[1,2]

[1] *Free University Amsterdam, and Tinbergen Institute,*
[2] *Netherlands Bureau for Economic Policy Analysis, The Hague.*

IDENTIFICATION OF 'WASTEFUL COMMUTING' USING SEARCH THEORY

Jos Van Ommeren

Free University
Amsterdam, the Netherlands, jommeren@feweb.vu.nl

Willemijn Van der Straaten

Free University
Amsterdam, the Netherlands, wstraaten@feweb.vu.nl
Netherlands Bureau for Economic Policy Analysis
The Hague, The Netherlands

22-09-2005

IDENTIFICATION OF 'WASTEFUL COMMUTING'
USING SEARCH THEORY

Abstract

In this paper, we employ search theory as a micro-economic foundation for the wasteful commuting hypothesis. It is argued that the commute of the self-employed is the result of a search process for vacant workplaces, whereas employees search for vacant jobs through space. Because the arrival rate of workplaces is much higher than the arrival rate of jobs, the self-employed essentially may minimise the commute, whereas employees accept jobs with a longer commute. In the empirical analysis, the extent of the 'wasteful commuting' is identified by estimating the difference in the commute of employees and self-employed individuals with fixed workplaces. Our estimates indicate that about 40 to 50% of the observed commute may be considered 'wasteful' due to job search imperfections. We reject alternative hypotheses why the self-employed have a shorter commute (self-selection of not working from home, different residence locations). In line with the theoretical model, the excess commute is shown to be less in areas with a higher urban density.

1. Introduction

The 'wasteful commuting' literature tests the assumption that workers optimally choose their residence or workplace location so that the costs related to commuting are minimised. Given the existence of a simplified static world with perfect labour or housing market markets, such an assumption would be plausible. In an economy with imperfections, however, this will not be the case (Weinberg et al., 1981; Zax, 1991; Holzer, 1994). In the wasteful commuting literature, market imperfections are defined as the presence of job and residential moving costs and lack of perfect information about job opportunities and vacant residences. Moving costs are relevant because they prevent employees to move job or workplace to reduce the commute, because the discounted moving costs exceed the benefit of a reduced commute. Imperfect information implies that employed (and unemployed) workers decide to accept jobs and residences, which do not minimise the commuting costs, because they do not have full information about all jobs and residences and have to search for vacant jobs and residences. The theoretical literature suggests that the length of the commute is increased by market imperfections. For example, Crane (1996) shows that uncertainty concerning job locations combined with positive residential moving costs increases the ratio of actual-to-minimum commuting in urban areas.

Although the wasteful commuting literature started as a test of the monocentric urban model (Hamilton, 1982, 1989; White, 1988), it is nowadays used to test the minimizing commuting costs assumption. A large number of studies use data on (intraregional and interregional) commuting flows to test this assumption.[1] For example, Cropper and Gordon (1991), Small and Song (1992), Manning (2003), Rodriguez (2004) and Kim (1995) provide evidence that more commuting occurs than the minimum amount required for workers to commute given the spatial distribution of job and residences. The best evidence based on regional commuting flows suggests that the ratio of actual-to-minimum commuting is at least two, so about 50% of the observed commute is 'wasteful' or 'excess'. Tests based on regional commuting flows are thought to be contaminated though, in particular because they rely on the strong assumption that workers are homogeneous with respect to jobs and residences (Kim, 1995; Manning, 2003).[2] Because jobs and residences are heterogenous, the calculated

---

[1] The standard procedure is to calculate the necessary minimum amount of commuting between regions if workers were assigned to job locations given the regional distribution of jobs and residences using a standard linear-programming problem.

[2] Studies such as Cropper and Gordon (1991) and Kim (1995) use additional micro information such as number of workers in household to add additional restrictions, but the number of additional restrictions are limited, so the criticism of homogeneity still applies.

minimum commuting distance must exceed the actual minimum distance.[3] Hence, it is often thought that taking heterogeneity into account that the excess commute is more likely to be *less* than 50%.[4][5]

In the current paper, we aim to employ an entirely different micro-economic approach. Our approach is to compare the commute of a representative group of workers with another group of workers for which it is plausible that the excess commute is negligible. We will argue that the excess commute must be much larger for *employees* than for the *self-employed*. In fact, we will argue that for the *self-employed* excess commute is more plausible to be absent.

This paper starts from the basis that the lack of full information about job opportunities implies that *employees* who search for jobs are confronted with a spatial distribution of acceptable job opportunities. Reformulating a standard job search model into a commuting model, we see that if the arrival rate of jobs is finite then the average commute exceeds the minimum commute, in line with the 'wasteful commuting' hypothesis. We argue further that the *self-employed* search for vacant workplaces (and not for jobs), and the arrival rate of workplaces is essentially infinite relative to the arrival rate of jobs, so the excess commute for the self-employed should be negligible. We show that the results are robust if we relax some of the restrictive assumptions. In addition, we show how the excess commute depends on spatial structure.

The aim of this paper is to estimate the extent of 'wasteful commuting' employing micro-economic data about the length of the commute of employees and self-employed workers. For both groups of workers, we use observations for which hold that the workers commute to their *fixed* workplace *on a daily basis* (and therefore do not work from home). We use two measures for the length of the commute – distance and time – and show that for both

---

[3] This issue is *partially* addressed by focusing on commuting flows of subgroups (e.g. different age classes), but this is unlikely to capture even a small part of the heterogeneity. Maybe surprisingly, studies such as Manning (2003) find that the excess commute using subgroups is close to using the aggregate group, probably because subgroups tend to work and live at the same locations as other subgroups.

[4] Rodriguez (2004) addresses the heterogeneity issue and focuses on a micro sample of bank tellers who work at different branches of the *same* firm within the same city. Hence, this study reduces the heterogeneity by focusing on a very specific subgroup with (almost) identical employment conditions except for location. In this study it is also reported that the excess commute is about 50%. The main disadvantage of focussing on excess commuting *within* a homogeneous group of workers is that the results are difficult to generalize. Note that in the study by Rodriguez (2004) it is ignored that bank tellers may reduce their commute by finding employment at *other* banks. Further, it is unclear to what extent these bank tellers may move residence to locations not occupied by a bank teller.

[5] Another difficulty with the usual tests is the choice of the size of the region, because the excess commute depends negatively on the size of the region (Small and Song, 1992), whereas the optimal size of the region is at least theoretically zero. Hence, by definition, regions are not as small as might be theoretically required.

measures, the excess commute is about 40 to 50% of the observed commute. The implications of the search model regarding urban density are confirmed.

One may argue that the methodology proposed is suspicious, because employees and self-employed operate in different (labour) markets and have many different household and location characteristics, which are difficult to observe. We deal with this criticism by using the Dutch 1998 labour force survey which contains many control variables related to sector and profession as well as multiple observations for each household. The results appear to be robust when we control for selection effects of not working from home as well as when we include household fixed effects to control for housing, location and household characteristics.

The outline of the paper is as follows. In Sections 2 and 3, we introduce a basic search model for employees and self-employed. In Sections 4 and 5, we will show the empirical results and in Section 6 we conclude.

2. A commuting model for employees

2.1 The basic model

A key assumption in the 'wasteful commuting' literature is that matches in the labour market and housing market are not unique. In other words, employees can be replaced by other employees who are equally productive and receive the same wage. We follow this literature by assuming that all jobs are identical and pay the same wage *w* but differ *only* with respect to the distance to the residence location, and therefore with respect to the commuting costs. The commuting costs are proportional to the commuting distance *t* and can be written as ηt (see also Manning, 2003). A priori, all workers are identical and face the same environment.

The assumption that all workers face the same environment essentially implies the absence of spatial structure and therefore the absence of house prices in the economy, in contrast to standard urban economic models which presume that the commute is fully compensated by lower house prices. The model used here may be interpreted as a model where house prices only weakly compensate for the commute, in line with most empirical results. More importantly, in the empirical analysis, we will estimate household fixed effects models, so we *fully* control for house attributes including the residence location and house price. By construction, workers within the same household face the same spatial structure.

Each worker is either unemployed (state 0) or employed as an employee (state 1). At random time intervals, an unemployed receives job offers randomly from each point in space at a rate λ. We initially assume that employees do not receive job offers. The commuting distance implied by a job offer is assumed to be a realisation of a random draw from a

cumulative employment density function $F(t)$, where $F(t)$ is the proportion of vacancies (employment offers) at a commuting distance no greater than t. We initially assume that $F(t)$ is continuous differentiable and strictly increasing in t. The unemployed accept or reject job offers as soon as they arrive. Given this set up, the unemployed accept jobs with a certain range, defined by the maximum acceptable commuting distance T (see, similarly, Van den Berg and Gorter, 1997). We initially assume that workers do not move residence.

We assume that employees are dismissed and thus become unemployed at rate δ. Any unemployed worker receives utility flow b per instant (b can be interpreted as an unemployment benefit). All individuals discount future income at rate r. Given the above assumptions, the expected discounted lifetime income when an individual is unemployed, V_0, can be expressed as the solution to the following Bellman equation:[6]

$$rV_0 = b + \lambda[\int \max\{V_0, V_1(x)\}dF(x) - V_0].$$

(1)

In words, lifetime income is equal to the flow of income while unemployed (the benefit) plus the expected gain in income attributable to finding acceptable jobs, where acceptance only occurs if the value of employment $V_1(t)$, exceeds that of continued search V_0. Similarly, the expected lifetime income of an employee who travels commuting distance t solves:

$$rV_1(t) = w - \eta t + \delta[V_0 - V_1(t)].$$

(2)

$V_1(t)$ is decreasing in t, whereas V_0 is independent of it, which implies that there exists a maximum acceptable commuting distance T, such that: $V_1(t) < V_0$ as $t > T$ and $V_1(t) > V_0$ as $t < T$.

Derivation of T is straightforward, since T is defined by $V_1(T) = V_0$. Equation (1) can be written as:

$$rV_0 = b + \lambda \int_0^T [V_1(x) - V_0]dF(x),$$

(3)

---

[6] Readers familiar with the labour market literature will recognize that the model presented is a standard search model, the only difference being that in the current model workers face a distribution of commuting distances, whereas the standard assumption is that workers face a wage distribution.

and, therefore equations (2) and (3) imply, using integration by parts, that:

$$T = \frac{w-b}{\eta} - \frac{\lambda}{r+\delta}\int_0^T F(x)dx.$$ (4)

It can be easily seen that if λ becomes larger, then the maximum acceptable commuting distance T is reduced, so the unemployed worker becomes more choosy and more job offers are rejected. In the extreme case that λ approaches infinity, T approaches zero. According to this reformulated standard search model, the employee's commuting distance t is positive (and less than T) due to search frictions because the arrival rate of jobs λ is finite. In equation (4), the employment density $F(t)$ is proportional to the arrival rate λ.[7] We will focus therefore on differences in λ, keeping $F(t)$ constant. The expected average commute is equal to $E(t|t < T)$.[8] It can be shown that $E(t|t < T)$ is decreasing in λ, because $E(t|t < T)$ is decreasing in T, and T is decreasing in λ (see (4)).

Let us presume now the absence of search frictions, so λ is infinite. So, the employee will effectively be able to choose the *optimal* commuting distance. In the current model, the optimally chosen commuting distance is the minimum distance, which is equal to zero. This can be easily seen, because V_1 is decreasing in t and F is differentiable everywhere. Hence, in the current model, a positive commuting distance t can be interpreted as 'wasteful commuting' and the average wasteful commute is equal to $E(t|t < T)$. Given the spatial distribution of jobs $F(t)$, T is defined by (4), so $E(t|t < T)$ is defined.[9] This result also holds when allowing for on-the-job mobility.[10]

---

[7] Hence a change in the arrival rate λ or a change in the employment density $F(t)$ has the same effect on T.

[8] $E(t|t<T) = \int_0^T [1-F(t|t \leq T)]dt$. In the current model $F(t|t \leq T) = F(t)/F(T)$, but this is not the case when we allow for on-the-job mobility.

[9] For example, suppose that firms are distributed homogenously over two-dimensional space, so $F(t) = \alpha t^2$, then $E(t|t < T) = 2/3T$.

[10] We have presumed above that on-the-job mobility is absent. Let us extend the basic model introduced above by allowing for on-the-job mobility. Jobs are assumed to arrive to the employed with arrival rate $\lambda_1 > 0$. Jobs are accepted by employees when the commuting distance is reduced. It can then be shown that (following the methodology by Burdett and Mortensen (1998)): $F(t|t < T) = (F(t)+\lambda_1 F(t)\delta)/(1+\lambda_1 F(t)/\delta)$. Hence $F(t|t < T)$ is increasing in λ_1/δ. When λ_1/δ approaches infinity, $F(t|t < T)$ approaches one for all $t \geq 0$, so $E(t|t \leq T) = 0$. Note that in most markets λ_1 is of the same magnitude as δ, hence this condition is never met in reality. Consequently, given on-the-job mobility, the job search model is still consistent with excess commuting as long as on-the-job mobility rates are finite.

The search model indicates that excess commute depends on the job arrival rate which may *not* be the same for all individuals. By assumption, for more specialized jobs, the mean job arrival rate (within a certain area) is less than for less specialized jobs. For example, the arrival rate for cleaning jobs within one kilometre from the residence must be much less than the job arrival rate for plumbing. Because specialization is strongly related to the degree of education, we hypothesize a positive relationship between the worker's excess commute and the worker's degree of education.

2.2 Relaxing assumptions on F(t)

We have assumed above that *F(t)* is continuous and strictly increasing in *t* ($t \geq 0$). This is not very realistic. One of the consequences of this assumption is that the minimum, and therefore optimal, commuting distance is equal to zero. We will relax this assumption now by assuming that *F(t) = H(t − τ)* for $t \geq \tau$ and *F(t) = 0* for $t < \tau$, where H is strictly increasing in its argument. In this case, τ is the minimum commuting distance (the distance to the nearest job) and $t - \tau$ is the extent of wasteful commuting. Hence, the average excess commute is equal to $E(t|t < T) - \tau > 0$.[11] Hence, given more realistic assumptions on F(t), the minimum commute τ is positive, and the average excess commute can be calculated as the average commute minus the minimum commute. In our empirical application, we will follow such an approach.

2.3 Residential mobility

We have presumed that residential mobility is zero, so, implicitly it is assumed that the residential moving costs are infinite. Let us presume now that the moving costs are finite, so workers will move residence when the commuting costs exceed a certain optimally chosen threshold T*. For simplicity, suppose that the location of the residence is optimally chosen by the employee and will be equal to zero. It can be shown that *T* = m / (r + δ)*, where *m* are the moving costs. T* is therefore equal to the discounted moving costs. Clearly, as long as *m > 0*, then T* > 0, and *F(t|t < T*) = F(t)/F(T*)*, so *E(t|t < T*) > 0* and the excess commuting is positive for employees.[12] This underlines that the excess commute is the combination of uncertainty about the location of jobs *and* positive residential moving costs.

---

[11] Let us give an example. Suppose that the nearest vacant place is at distance τ. Further, we presume that space is homogenous and two-dimensional, so $H(t) = \alpha t^2$. It can then be easily shown that $E(t|t < T) = \tau + 2/3(T-\tau)$, and $2/3(T-\tau)$ is the average size of the wasteful commuting. In the specific case that $T = 2\tau$, then the average excess commute is equal to $2/3\tau$, so the excess commute as a share of the average commuting distance is equal to 40%.

[12] In case T* exceeds T then the worker would never move residence, so the moving costs are essentially infinite.

2.4. Heterogeneous spatial structure

In the previous sections we ignored heterogeneous spatial structure as we presumed that all employees face the same spatial structure $F(t)$. This assumption is in contrast to the idea, essential in the wasteful commuting literature, that employees commute to other regions, because employment and population are not evenly distributed over space. In empirical studies, it is usually found that employment is more spatially concentrated than population. Suppose therefore the existence of two regions of equal size in terms of population, but not in terms of employment. Of the total employment, the share of employment in region 1 is equal to $\sigma \leq \frac{1}{2}$ and in region 2 the employment share is $1-\sigma$. Suppose that the regions are at a distance of each other such that the interregional costs are less than T and that the distances within the region are negligible. Hence, we have essentially rewritten the job search model in a way that job offers are always accepted in both regions, but workers prefer to receive a job offer from their region of residence. We standardize the distance between the two regions and the total size of the population to one.

Assuming the absence of search imperfections, the minimum average commuting costs in the two-region economy can easily be seen to be equal to $\eta(\frac{1}{2}-\sigma)$. Hence, the minimum commute depends negatively on σ which measures the distribution of employment relative to population over space. Only when $\sigma = \frac{1}{2}$, so employment is homogeneously distributed over space, then the minimum average commute is zero (in line with the assumption of the basic model of Section 1). Now suppose the presence of search imperfections. Presume that the probability of receiving a job offer in a region is proportional to the size of the region.[13] In this case, the average commuting costs can easily be shown to be equal to $\frac{1}{2}\eta$. Consequently, the excess commute due to search perfections, measured by the average commute minus the minimum average commute, is equal to $\sigma\eta$. The implication is that the *excess commute due to search imperfections is increasing in* σ, indicating that the excess commute is *less* in urban economies where employment is spatially concentrated, such as is common in strongly agglomerated areas with high employment densities.[14] This result is intuitive: the excess commute is the result of accepting jobs that do not minimize commuting costs due to *spatial* uncertainty about future job offers. This uncertainty depends negatively on the spatial

---

[13] This assumption seems a reasonable approximation and can be derived from the assumption that job seekers search randomly over space *and* the number of regional vacancies is proportional to the number of regional jobs.
[14] Note that the empirical literature tends to focus on these areas, suggesting that empirical studies underestimate the average excess commute.

employment concentration.[15] The excess commute obtains it's maximum when employment is relatively homogeneously distributed over space, as is more common in rural areas with low employment densities.[16] Hence, the testable implication is that the excess commute depends negatively on employment density.

3. A commuting model for the self-employed

In the previous section, we have focused on employees. We will focus now on the self-employed. The self-employed search for suitable workplaces, not for jobs. Hence, we analyse the search process of a self-employed who is looking for a suitable vacant workplace location for her company. So, we can use a similar search model as used for the employee (see Section 2), the main difference being that the self-employed searches for a vacant workplace (and not for a job). Another distinction between employees and self-employed is that a large proportion of the self-employed work from home. The decision to work from home is plausibly exogenous with respect to the search process for suitable workplaces.[17] This decision is determined by factors such as the suitability of the residence as a workplace (restrictions due to type of work, the presence of subordinates). We will concentrate here on self-employed who do not work from home.[18]

The main distinction between the search process of employees and self-employed is that the arrival rate of suitable workplaces is *much* higher than the arrival rate of suitable jobs. The main reason is that the job searcher looks for a job which matches her skills. The density of workplaces is much higher than the density of jobs that match her skills.[19] Hence, from a search perspective, the workplace arrival rate is close to infinite, so the excess commute due to search frictions should largely disappear (if λ approaches infinity, then the excess commute is zero). Hence, the difference between the commute of employees and self-employed can be

---

[15] In the extreme case that employment is concentrated in one point, then each worker knows where the (new) job will be located.

[16] In the case of isolated small settlements, the excess commute will also be small, because there is "no option to commute somewhere else". This case will be relevant for example for the US, but not for countries such as the Netherlands where such settlements are absent, and where employment density is relatively high everywhere.

[17] It seems less likely that working from home is the outcome of an unsuccessful search process for suitable workplaces, where workplaces are not suitable because the commuting length to the workplace is considered to be too long.

[18] Note that in the empirical analysis, we will select only workers who do not work from home and we take this selection effect into account.

[19] For example the spatial density of suitable workplaces may easily be thousand times larger than those of jobs for any occupation which require a standard office. Other reasons are that the job searcher has to be accepted for the job by the employer and that office vacancy rates tend to exceed job vacancy rates.

interpreted as the excess commute due to job-search imperfections.[20] This result also holds allowing for changes in the workplace location of the firm of the self-employed and allowing for residential mobility. The results of Section 2.4 on heterogeneous spatial structure indicate that if the self-employed systematically locate their residence in other locations than the employees, then our measure of the excess commute is flawed. We will show however empirically by employing household fixed effects that this is *not* the case.

4. Descriptive data

In this paper, we use the Dutch labour force survey (1998). It distinguishes between those who commute and those who work from home. Further, and this is essential for our application, it contains information about whether the worker has commuted on a daily basis to a *fixed* workplace location (for a period of at least three months before the interview date). It also includes essential background information such as the number of working hours, size of firm and detailed information about sector and occupation. In total, we analyse 44,260 observations of workers with a fixed workplace. In the whole sample, 11% is self-employed.[21] In Tables 1 and 2, we give the basic descriptives of the main variables of interest (descriptives of other explanatory variables can be found in Appendix 1). As can be seen from Table 1, the self-employed are much more likely to work from home than the employees (47.3 % versus 0.8 %), and their average commuting time is much smaller (14.4 minutes versus 22.5 minutes for workers who do not work from home).[22] Further, they have a shorter commuting distance (see Table 2). The preferred measure of the length of the commute is commuting distance, because commuting time is influenced by the endogenously chosen speed which may differ between employees and self-employed.[23] Nevertheless, both measures have their merits, so

---

[20] Another implication is that employees are more likely than the self-employed to move residence to decrease the commute. It implies for example that the probability to search for another residence depends positively on the commuting time for employees, but this effect is less strong for the self-employed.

[21] For a few observations the distinction between employees and self-employed is ambiguous. Self-employed who work for a firm owned by their spouse or parents are therefore excluded in the analysis.

[22] In the survey, we do not have information about the exact commuting time if it exceeds 60 minutes. This occurs in 7% of the observations. We initially presume that the commuting time for these observations is 60 minutes.

[23] One may argue that workers search over space in terms of distance and then choose the optimal commuting speed endogenously. The self-employed are more likely to use a car, which may result in a higher speed and therefore a smaller commuting time which biases the measure of excess commute. We will show later on that this is the case. On the other hand, one may argue that time is a more accurate measure of costs of travel, because of differences in speed levels available to workers, e.g. due to differences in road networks, congestion etc. Another argument is that time is usually more precisely reported by workers. Hence, measurement errors are likely to be smaller for commuting time as a proxy for commuting costs. For commuting distance, the measurement error is likely larger, but less likely systematically different for employees and self-employed. Because measurement error in the dependent variable does *not* affect the estimated effects, commuting distance is the preferred measure.

we will use both measures. It appears that both measures render similar estimates of the average excess commute. The descriptive data suggest that the excess commute is on average 36%, measured in minutes, and about 31% measured in distance.[24]

5. Results

5.1 Main results

We have selected those workers who do not work from home, commute on a daily basis, have a fixed workplace and have a positive commuting time and distance.[25] This guarantees that the commute of employees and self-employed are comparable. The endogenous variables are the logarithm of commuting time and distance. We include a large number of explanatory variables including size of the firm, address density, educational level, gender, occupation, industry, age, presence of children, number of hours worked, and presence of spouse and some interactions with gender.[26] We employ OLS for commuting time.[27] We estimate the effect on commuting distance, taking into account that distance is reported in classes (interval regression). The empirical results for commuting time and distance can be found in Tables 3 and 4.[28]

On average, the commuting time and distance of the self-employed is considerably less than those of the employees. This result is in line with the theoretical model, which predicts that the self-employed have shorter commutes, because they are not confronted with job search imperfections. Accordingly, the extent of the excess commute is equal to the difference in the average commute of employees and self-employed. Based on the results of columns 1 of Tables 3 and 4, about 40 to 50% of the commute of workers is due to job search

---

[24] The average commuting distance is calculated by taking the mid-points for all categories, and presuming 50 km for the last category. The average commuting distance is then 15.4 km for employees (in line with other studies for the Netherlands) and 10.6 km for the self-employed.

[25] For a few workers, reported commuting time is zero although they state that they do not work from home. These observations are deleted.

[26] We do not have information about income. It has been shown that ceteris paribus, the self-employed earn more (see e.g. Taylor, 1996). Further, it is generally found that there is a weak, but *positive* association between income and commuting time (see e.g. Manning, 2003). Consequently, our estimates are a slight underestimate of the extent of the excess commute. Note that Rouwendal and Rietveld (1994) include income and do not find any effect for the Netherlands, so the potential bias is likely negligible.

[27] We have re-estimated the effect on commuting time using Tobit models taking into account that for some observations we observe only that it exceeds 60 minutes. The results are identical.

[28] A large number of micro-economic studies is concerned with the empirical analysis of commuting distance or time (e.g. White, 1986; Rouwendal and Rietveld, 1994; Benito and Oswald, 1999; Van Ommeren et al., 1999). A notable result of the current as well as these studies is that the reported R^2 is typically low. This suggests that commuting is mainly an outcome of a stochastic process in which the lack of information plays an important role.

imperfections, and therefore 'wasteful'.[29] We consider this to be our main basic result. This result seems to be in line with most studies based on aggregate data. This suggests that the heterogeneity issue is of less importance than usually thought in the literature, for example because heterogeneous workers tend to work at the same workplace locations and live in the same neighbourhoods (as shown by Manning (2003)). We will now continue by examining extensions as suggested by theory.[30]

5.2 Excess commute and specialization

According to the job search model, the excess commute depends negatively on the arrival rate of jobs within a certain area and therefore depends negatively on the employment density of jobs. The higher educated search for more specialized jobs for which the employment density is much lower, so the excess commute should be higher for this subgroup. Therefore, we re-estimate models with the interaction of different levels of education and self-employed. The results can be found in the second columns of Tables 3 and 4. For commuting distance, the results are in line with the job search model. The self-employed with a university degree commute about about 40% less than self-employed with only a basic education.[31]

5.3 Urban density

In the theory, we have hypothesized that the excess commute may depend on the employment density. We examine this hypothesis by including an interaction effect between self-employed and *address* density. Address density measures the average number of addresses per square kilometre (within one kilometre of each residence) for each municipality. Addresses may refer to residences, firms, schools, shops etc. Note that address density is a mixture of population and employment density, the latter being the ideal measure. We observe address density in classes. To avoid multicollinearity, we measure address density by means of a continuous variable which is equal to the logarithm of the mid-points of the address density classes. In line with the theoretical model, we find that the excess commute is larger in *less* urban areas. As can be see in column (3) of Table 2, the elasticity of address density on the excess commute is 0.109 (with a standard error of 0.025). This turns out to be a sizeable effect. For

---

[29] The difference in excess commute between measures based on distance and time is rather small, in line with Small and Song (1992).

[30] The effect of explanatory variables is in line with other studies in the literature (see e.g. Rouwendal and Rietveld, 1994). For example, also in line with the theoretical model, the commuting distance and time are much longer for workers with a higher educational degree.

[31] For commuting *time*, this particular prediction of the search model does not seem to hold, but we will see later on that when we include household fixed effects, then the prediction does hold.

example, when we use commuting time, the excess commute in areas with an address density of 2500 is about 30%, whereas in areas with an address density of 500, the excess commute is about 50%. When we use commuting distance, the same conclusion is drawn. The implied excess commute in Amsterdam, the largest city in the Netherlands (with a density of about 6000) is around 20%, so the estimates do *not* imply that the excess commute disappears in large urban areas.

5.4 Sensitivity analysis

5.4.1 Self-selection

One potential bias of our results is that the analysis is based on a sample of workers who do not work from home. Self-selection of working from home may bias our results.[32] We control therefore for the selection of working from home using a standard Heckman procedure. As additional instruments in the selection equation, we use *interactions* of being self-employed with a number of variables: gender, children, size of firm, working hours.[33] Correcting for self-selection appears to have no effect on the results: the results remain unchanged. Results can be received upon request from the authors.

5.4.2 Housing market controls

Our analysis is based on a labour force survey, so controls for labour market variables are abundantly available, but we essentially lack good controls related to the residence, and in particular the residence location of the household. It may be argued that the self-employed live systematically at different residence locations than the employees, so our measure of the excess commute is seriously biased. We therefore re-estimate the above models including household fixed effects models. This implies that *for each household* one dummy is included, hence the estimates are essentially based on the differences of the length of the commute of workers *within* the household only. It can be seen from the last two columns of Table 3 that the results remain essentially the same for commuting time. The only difference, as can be seen in the last column, is that, now in line with theory, it appears that the excess commute is

---

[32] For example, it may be the case that the self-employed are more likely to work from home when the minimum distance at which they can find a workplace is large.

[33] These instruments can be justified on the ground that working from home is common for the self-employed but not for the employees. Note that these additional instruments are not necessary in the case of commuting time because of identification via the functional form.

larger for the highly educated.[34] We lack enough variation in commuting distance to estimate the effect on distance using household fixed effects.[35]

5.4.3 Labour market controls

Many employees are in sectors in which the self-employed are not present, for example the public sector etc. Although we control for the type of sector, it may be the case that due to interactions with other explanatory variables our estimates of the effect of being self-employed on the commute are biased. We have therefore re-estimated the models selecting observations of workers employed in sectors in which at least 1% of the workers is self-employed. It appears that the results are essentially the same. The results can be received upon request from the authors.

We have also estimated additional models which allow for interaction effects of self-employed with other explanatory variables. In particular, the interaction effect of gender turns out to be relevant. A priori, it is not clear whether the excess commute should be larger, or smaller, for males than for females, because it depends on gender differences in the job arrival rate λ and the value of not working b. In particular, gender differences in the value of not working are considered to be large. Given the common assumption that the value of not working is higher for females due to alternative use of time in the household, the excess commute of males should exceed the excess commute of females. Our empirical results suggest that this may be the case: the excess commute for females is 27%, whereas for males it is 40% when using commuting time. Based on commuting distance, the females' excess commute is 30%, whereas the males' excess commute is 55%.

5.4.4 Differences between time and distance

Note that the reported extent of excess commuting using a measure based on commuting *time* is slightly higher then using a measure based on *distance*, suggesting that the average commuting *speed* of the self-employed is slightly higher than the average speed of employees. One reason maybe that the self-employed use the car for *short* distances, because they are more likely to need the car during work hours, whereas for longer distances, the choice in

---

[34] Remark that explanatory variables that are the same for workers within a household such as address density are not identified, but the interaction of address density with self-employed is identified.

[35] Recall that commuting distance is measured in four classes. The effect of being self-employed is identified based on observations of households where at least one employee and at least one worker are present *and* both the employee and worker report a commuting distance of different length. For other explanatory variables a similar condition holds. It appears that there is too little variation in the data to estimate the household fixed effects model for distance measured in classes.

mode between employees and self-employed is small. In Appendix 2, we show that the self-employed commute slightly faster - about 9% - when one controls for commuting distance.[36] Hence, plausibly, the measure of excess commuting based on commuting distance is more reliable, but, most likely, the potential bias of using time is small, in line with Small and Song (1992).

6. Conclusion

The assumption that labour markets are perfect has been frequently criticised (e.g. Anas, 1982; Hamilton, 1982, 1989). In particular, it may be argued that imperfect information about job opportunities are ignored. An essential characteristic of the labour market is therefore that individuals have to search for jobs so jobs are accepted that do not minimize the commuting costs, inducing an excess commute.

In this paper we have analysed the extent of 'wasteful commuting' using a micro-economic approach which is essentially based on a comparison of the commute of employees and self-employed individuals. We have argued that for the self-employed workers, the excess commute may be thought to be absent. Our main conclusion is that about 40 to 50% of the commute is due to job search imperfections, and therefore 'wasteful'. The results are in line with results reported in the literature based on aggregate commuting flows, which ignores heterogeneity among workers. This is a bit surprising because it is generally thought that heterogeneity causes an upward bias, which is potentially large (see Manning, 2003).

Our theoretical and empirical analyses further indicate that the excess commute is smaller in more urban areas, suggesting that studies such as White (1988) and Hamilton (1989) that focus on large urban areas may be downward biased. One tentative interpretation of our study is then that by focusing on urban areas, such studies underestimate the excess commute, but by ignoring heterogeneity, they overestimate the excess commute, so the overall bias may be not so large. Another interpretation is that heterogeneity is not so much a problem, as argued by Manning (2003), so aggregate and micro data analysis should find estimates of excess commute of similar magnitude.

---

[36] Note that the analysis of speed is structurally different from the separates analyses of commuting time and distance, because in the speed analysis, we control for commuting distance. This small effect is not surprising in the light that for *short* commuting distances the differences in speed between car and bicycle (the main alternative to the car in the Netherlands) is small. Van Ommeren and Dargay (2006) report for the UK even smaller differences between the average travel speeds of employees and the self-employed. Note that because distance is observed in classes, also speed is measured in classes. For the explanatory variable distance, we use the mid-point of the classes.

Reference

Altmann, J.L. and J.S. DeSalvo (1981), Test and extensions of the Mills-Muth simulation model of urban residential land use, *Journal of Regional Science*, 21, 1-21.

Anas, A. (1982), *Residential location markets and urban transportation: economic theory, econometrics and policy analysis with discrete choice models*, Academic Press, New York.

Benito, A. and A.J. Oswald (1999), Commuting in Great Britain in the 1990's. University of Warnick.

Blanchflower, D. and B. Meyer (1994), A longitudinal analysis of the young self-employed in Australia and the United States, *Small Business Economics*, 6, 1, 1-19.

Buliung, R. and P.S. Kanaroglou (2002), Commute minimization in the Greater Toronto Area: applying a modified excess commute, *Journal of Transport Geography*, 10, 177-186.

Burdett, K. and D.T. Mortensen (1998), Wage differentials, employer size, and unemployment, *International Economic Review*, 39, 2, 257-293.

Crane, R. (1996), The influence of uncertain job location on urban form and the journey to work, *Journal of Urban Economics*, 37, 342-56.

Cropper, M. & Gordon, P. (1991), Wasteful commuting: a re-examination, *Journal of Urban Economics*, 29, 2-13.

Hamilton, B.W. (1982), Wasteful commuting, *Journal of Political Economy*, 90, 1035-1053.

Hamilton, B.W. (1989), Wasteful commuting again, *Journal of Political Economy*, 97, 1497-1504.

Holzer, H. J. (1994), Work, search and travel among white and black youth, *Journal of Urban Economics*, 35, 320-345.

Kim, S. (1995), Excess commuting for two-worker households in the Los Angeles metropolitan area, *Journal of Urban Economics*, 38(2), 166-182.

Manning, A. (2003), The real thin theory: monopsony in modern labour markets, *Labour Economics*, 10, 105-131.

Rodriguez, D. (2004), Spatial choices and excess commuting: a case study of bank tellers in Bogota, Colombia, *Journal of Transport Geography*, 12, 49-61.

Rouwendal, J. and P. Rietveld (1994), Changes in commuting distances of Dutch households, *Urban Studies*, 31, 09, 1545-1557.

Small, K. & S. Song (1992), Wasteful commuting: a resolution. *Journal of Political Economy*, 100, 888-898.

Taylor, M. (1996), Earnings, independence or unemployment: why become self-employed?, *Oxford Bulletin of Economics and Statistics*, 58, 2, 253-266.

Van den Berg, G.J. and C. Gorter (1997), Job search and commuting time, *Journal of Business and Economic Statistics*, 15, 2, 269-281.

Van Ommeren, J.H., P. Rietveld and P. Nijkamp (1999), Job moving, residential moving and commuting: a search perspective, *Journal of Urban Economics*, 46, 230-253.

Van Ommeren J.H. and J. Dargay (2006), The optimal choice of commuting speed: consequences for commuting time, distance and costs, *Journal of Transport Economics and Policy*, forthcoming.

Weinberg, D.H., J. Friedman, and S.K. Mayo (1981), Intraurban residential mobility: the role of transaction costs, market imperfections and household disequilibrium, *Journal of Urban Economics*, 9, 332-348.

White, M. J. (1986), Sex differences in urban commuting patterns, *American Economic Review*, 76, 368-372.

White, M.J. (1988), Urban commuting journeys are not "wasteful", *Journal of Political Economy*, 96, 1097-1110.

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

Table 1: Means

	Employee	Self-employed
Work from home	0.8 %	47.3 %
Commuting time *	22.5 min.	14.4 min.
Commute by car *	57.8 %	70.8 %

Note: * only for those who work outside the home.

Table 2: Commuting distance

Commuting distance	Employee	Self-employed
1) 0 - 7 km	45.4 %	66.1 %
2) 8 - 17 km	25.0 %	17.6 %
3) 18 - 32 km	16.9 %	9.1 %
4) > 32 km	12.7 %	7.2 %

Note: only for those who work outside the home.

Table 3: Commuting time

	(1) log (time)	(2) log (time)	(3) log (time)	(4) log (time)	(5) log (time)
Self-employed	**-0.387** (0.030)			**-0.431** (0.065)	
Basic education * Self-employed		**-0.445** (0.088)	**-0.498** (0.089)		**-0.387** (0.183)
Lower secondary education * Self-employed		**-0.322** (0.055)	**-0.354** (0.053)		-0.184 (0.109)
Higher secondary education * Self-employed		**-0.416** (0.038)	**-0.441** (0.039)		**-0.402** (0.081)
Higher vocational education * Self-employed		**-0.324** (0.051)	**-0.379** (0.053)		**-0.422** (0.102)
University * Self-employed		**-0.429** (0.054)	**-0.487** (0.055)		**-0.703** (0.111)
Lower secondary education	-0.015 (0.019)	-0.020 (0.020)	-0.022 (0.020)	-0.024 (0.039)	-0.030 (0.040)
Higher secondary education	**0.052** (0.019)	**0.051** (0.019)	**0.050** (0.019)	0.044 (0.040)	0.046 (0.040)
Higher vocational education	**0.170** (0.022)	**0.165** (0.022)	**0.165** (0.022)	**0.100** (0.045)	**0.107** (0.046)
University	**0.300** (0.025)	**0.301** (0.025)	**0.312** (0.025)	**0.196** (0.055)	**0.223** (0.057)
Address density $\geq$ 2500 per km^2	**0.083** (0.016)	**0.083** (0.016)			
Address density 1500 - 2500 per km^2	-0.008 (0.013)	-0.009 (0.013)			
Address density 1000 - 1500 per km^2	-0.015 (0.013)	-0.015 (0.013)			
Address density 500 - 1000 per km^2	**-0.032** (0.013)	**-0.030** (0.013)			
Log (address density)			0.011 (0.006)		
Self-employed * log (address density)			**0.109** (0.025)	**0.190** (0.055)	**0.205** (0.056)
Male	0.017 (0.019)	0.017 (0.019)	0.013 (0.019)	**-0.130** (0.057)	**-0.128** (0.057)
Couple	**0.047** (0.015)	**0.047** (0.015)	0.043 (0.015)	0.041 (0.058)	0.041 (0.058)
Children < 18 years	**0.035** (0.012)	**0.035** (0.012)	0.032 (0.012)		
Couple* Female	-0.020 (0.022)	-0.021 (0.022)	-0.023 (0.022)	**-0.189** (0.057)	**-0.188** (0.057)
Child * Female	**-0.061** (0.018)	**-0.061** (0.019)	**-0.061** (0.019)	-0.040 (0.024)	-0.040 (0.024)
Age < 30	0.027 (0.015)	0.027 (0.015)	0.028 (0.015)	**0.113** (0.052)	**0.115** (0.052)
Age 30- 50	**0.027** (0.013)	**0.027** (0.013)	0.028 (0.013)	0.033 (0.038)	0.032 (0.038)
Workhours/10	**0.200** (0.018)	**0.198** (0.018)	**0.198** (0.018)	**0.234** (0.032)	**0.229** (0.032)
Workhours^2/1000	**-0.233** (0.024)	**-0.230** (0.024)	**-0.230** (0.024)	**-0.263** (0.043)	**-0.254** (0.044)
Size firm < 10	**-0.279** (0.016)	**-0.279** (0.016)	**-0.282** (0.016)	**-0.262** (0.029)	**-0.261** (0.030)
Size firm 10-99	**-0.177** (0.010)	**-0.177** (0.010)	**-0.178** (0.010)	**-0.119** (0.020)	**-0.117** (0.020)
17 sectors	Included	Included	Included	Included	Included
Occupations (83)	Included	Included	Included	Included	Included
Household fixed effects				Included	Included
R^2	0.115	0.115	0.113	0.708	0.708
Standard error	0.773	0.773	0.773	0.700	0.700
Number of observations	34634	34634	34634	21612	21612

The references are: employees, basic education, address density < 500 per km^2, female, single, no children, age $\geq$ 50, size firm > 99. The number of observations of the household fixed effects model is the number of households with at least two workers.

Table 4: Commuting distance

	(1) log (distance)	(2) log (distance)	(3) log (distance)
Self-employed	**-0.494**		
	(0.052)		
Basic education * Self-employed		**-0.341**	**-0.423**
		(0.169)	(0.170)
Lower secondary educ * Self-employed		**-0.284**	**-0.333**
		(0.098)	(0.099)
Higher secondary educ * Self-employed		**-0.454**	**-0.493**
		(0.067)	(0.068)
Higher vocational educ * Self-employed		**-0.492**	**-0.577**
		(0.089)	(0.091)
University * Self-employed		**-0.725**	**-0.804**
		(0.093)	(0.096)
Lower secondary educ	-0.011	-0.013	-0.014
	(0.034)	(0.034)	(0.034)
Higher secondary educ	**0.156**	**0.161**	**0.161**
	(0.033)	(0.033)	(0.033)
Higher vocational educ	**0.370**	**0.380**	**0.379**
	(0.037)	(0.038)	(0.038)
University	**0.567**	**0.594**	**0.595**
	(0.042)	(0.043)	(0.043)
Address density $\geq$ 2500 per km^2	**-0.482**	**-0.484**	
	(0.027)	(0.027)	
Address density 1500 - 2500 per km^2	**-0.411**	**-0.412**	
	(0.022)	(0.022)	
Address density 1000 - 1500 per km^2	**-0.301**	**-0.302**	
	(0.022)	(0.022)	
Address density 500 - 1000 per km^2	**-0.152**	**-0.153**	
	(0.022)	(0.022)	
Log (address density)			**-0.209**
			(0.009)
Self-employed * log (address density)			**0.169**
			(0.044)
Male	**0.156**	**0.158**	**0.156**
	(0.032)	(0.032)	(0.032)
Couple	**0.116**	**0.116**	**0.118**
	(0.026)	(0.026)	(0.026)
Children < 18 years	-0.016	-0.014	-0.014
	(0.021)	(0.021)	(0.021)
Couple* Female	0.041	0.042	0.040
	(0.037)	(0.037)	(0.037)
Child * Female	-0.030	-0.031	-0.030
	(0.031)	(0.031)	(0.031)
Age < 30	**0.147**	**0.146**	**0.147**
	(0.025)	(0.025)	(0.025)
Age 30- 50	**0.150**	**0.149**	**0.150**
	(0.022)	(0.022)	(0.022)
Workhours/10	**0.358**	**0.359**	**0.356**
	(0.032)	(0.032)	(0.032)
Workhours^2/1000	**-0.372**	**-0.373**	**-0.370**
	(0.042)	(0.042)	(0.042)
Size firm < 10	**-0.395**	**-0.392**	**-0.393**
	(0.027)	(0.027)	(0.027)
Size firm 10-99	**-0.242**	**-0.240**	**-0.239**
	(0.018)	(0.018)	(0.018)
17 sectors	Included	Included	Included
Occupations (83)	Included	Included	Included
Log Likelihood	-40766	-40760	-40759
N	34004	34004	34004

The references are: employees, female, basic education, single, no children, age $\geq$ 50, size firm > 99, address density < 500 per km^2.

Appendix 1: Descriptives

	Mean	
	Employed	Self-employed
Basic educ.	0.066	0.053
Lower sec. educ	0.199	0.163
Higher sec. educ	0.448	0.461
Higher voc. educ	0.195	0.170
University	0.092	0.153
Address density $\geq$ 2500 per km^2	0.172	0.167
Address density 1500 - 2500 per km^2	0.253	0.222
Address density 1000 - 1500 per km^2	0.220	0.217
Address density 500 - 1000 per km^2	0.205	0.227
Address density $\leq$ 500 per km^2	0.151	0.168
Male	0.579	0.789
Single	0.160	0.123
Couple	0.738	0.850
Children < 18 years	0.531	0.547
Age < 30	0.283	0.084
Age 30 - 50	0.559	0.616
Age > 50	0.058	0.147
Working hours	34.44	51.81
Size firm < 10	0.100	0.812
Size firm 10 -99	0.272	0.174
Size firm > 99	0.628	0.014

Appendix 2: Speed

	(1) ln (speed)
Ln (commuting distance)	**0.788**
	(0.006)
Self-employed	**0.090**
	(0.025)
Lower secondary education	0.005
	(0.015)
Higher secondary education	0.015
	(0.015)
Higher vocational education	-0.016
	(0.017)
University	**-0.068**
	(0.019)
Address density $\geq$ 2500 per km^2	**-0.245**
	(0.118)
Address density 1500 - 2500 per km^2	**-0.154**
	(0.010)
Address density 1000 - 1500 per km^2	**-0.098**
	(0.009)
Address density 500 - 1000 per km^2	**-0.042**
	(0.009)
Male	**0.034**
	(0.015)
Couple	**0.038**
	(0.012)
Children < 18 years	**-0.044**
	(0.009)
Couple* Female	0.028
	(0.017)
Child * Female	**0.043**
	(0.014)
Age < 30	0.013
	(0.011)
Age 30- 50	**0.032**
	(0.010)
Workhours/10	-0.014
	(0.015)
Workhours^2/1000	0.037
	(0.020)
Size firm < 10	**0.055**
	(0.012)
Size firm 10-99	**0.046**
	(0.008)
17 sectors	Included
Occupations (83)	Included
Log Likelihood	-10880
N	33902

Note: The explanatory variable Ln(commuting distance) is the midpoint of the commuting distance class as reported in Table 2. Estimates are based on an interval regression because speed is observed in classes.