The potential effects of transport costs on unemployment in South Africa

Kieran de Lange

Declaration

I declare that this dissertation is my own unaided work. It is submitted for the Master of Economics degree (School of Economic and Business Sciences) at the University of the Witwatersrand, Johannesburg. It has not been submitted for any other degree or examination at any other University. I also declare that this is my original work produced and submitted in accordance with all rules of professional academic standards and ethics.

_____Signature of candidate

_____Day of_____2013

Acknowledgements

I would like to acknowledge the following people for their contribution to this work:

1. My supervisor, Dr Neil Rankin, for his valuable contributions to this paper, as well his belief that I could complete my Master's degree. Without him, I would not be here.
2. Gareth Roberts, for his assistance with Stata.
3. My family and friends, for their continued support.
4. Claire Sebastiao, because she's a legend.

Table of Contents

(i) Abstract

The relationship between transport costs and unemployment in South Africa is under researched, mostly due to a lack of appropriate data. However, two surveys—the National Income Dynamics Study (NIDS) and the Labour Market Entry Survey (LMES)—have recently collected information on transport costs. This thesis uses these new sources of data to examine three aspects of the relationship between transport costs and unemployment in South Africa. The first is the relationship between transport costs and the probability of searching for employment. The second is the relationship between reservation wages and transport costs. The third is the relationship between transport premiums and transport costs, with a focus on commute times as a proxy for transport costs. This research shows: the negative correlation between job search and transport costs; a positive correlation between median transports costs and reservations wages for jobs that are far; negative correlation between transport premiums and transport costs.

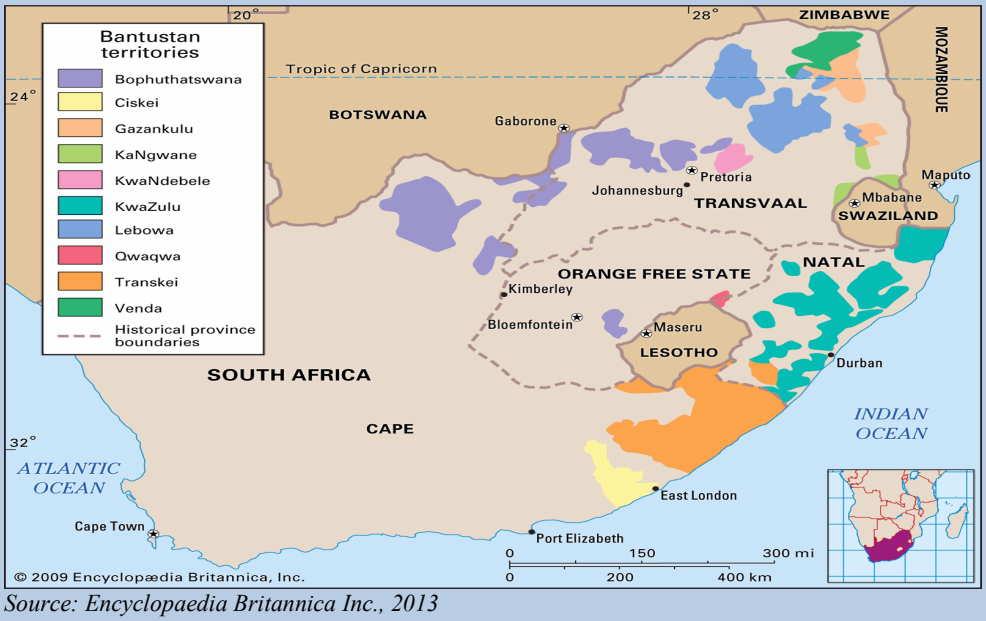
(ii) Introduction

South Africa has an unemployment problem. Since 1994, the number of unemployed South Africans has consistently increased. A key contributing factor has been the entry into the labour market of those who were previously excluded from it. With a large pool of previously disadvantaged and jobless Africans suddenly entering the labour market in 1994, the new government faced the monumental task of creating jobs. They were partially successful, but not to the extent required by the economy: 1.6 million jobs were created from 1995 to 2002, but 5 million job seekers entered the job market during the same period (Bhorat, 2004). Kingdon and Knight (2006) report that the labour force grew by 2.6% and 2% annually during the 1980s and 1990s, respectively. During the same periods, formal wage employment declined by 1.2% and 1.6%, respectively (Kingdon and Knight, 2006).

Since 1994, although overall employment numbers have increased, the unemployment rate has remained persistently high. Several studies have investigated the South African unemployment puzzle. Cichello et al. (2005) assert that one of the main reasons for this unemployment rate is the small number of opportunities available to the numerous job seekers in the South African labour market. Nattrass (2000a), however, argues that wage rigidity—as a result of labour legislation—hinders the labour market from creating formal-sector jobs, because those who are currently employed are reluctant to take home lower wages. This leaves fewer options available to employers. Arora and Ricci (2005) and Kingdon and Knight (2006) also point to labour market regulation as a significant obstacle to increased employment, while Arora and Ricci (2005) indicate that centralised bargaining and the role of unions, with regard to labour standards, are a major force in wage stickiness and unemployment. These factors undoubtedly contribute to unemployment, but there has been much less research on the effect that spatial inequality and the resulting high transport costs to and from work have on unemployment.

The legacy of Apartheid in South Africa is an unemployment problem that is geographical or spatial in nature. The spatial mismatch hypothesis (SMH), which was first described by Kain in 1968, proposes that marginalised African American workers live in segregated zones (city centres) that are separated from major employment areas (suburbs). The SMH is particularly relevant to South Africa

with its history of legislated segregation. The Group Areas Act of 1950 certainly contributed to this spatial mismatch. (Louw, 2010). The Act was initiated in 1954 and it segregated people based on the colour of their skin and also segregated amenities (Mabin, 1992). The Group Areas Act formally included Coloured people in the retroactive deprivation of tenurial rights and physical removal from occupied mixed areas (Mabin, 1992). The Act identified four groups: White, Indian, Coloured and African. Coloured and Indian people were often relocated to new suburbs away from the inner city, resulting in them being far from their places of work (Louw, 2010). Africans were forcibly moved to areas known as locations or model locations (Bantustans), where they faced similar problems to those faced by Indians and Coloureds in terms of distance to work. The Bantustans, created through the 1950 Act, were located far from main city centres, which meant access to employment required longer commutes and more money. The image below shows the distribution of Bantustans created through the 1950 Act. The main cities are marked with stars, clearly showing that the group areas were located some distance away from the cities, which meant that access to employment required longer commutes and more money. Therefore, a key factor that may contribute to high reservation wages is that people may face high transport costs commuting to and from job locations. This means that net wages may be much smaller and may account for findings such as Walker's (2003), who finds that 15% of the Khayelitsha-Mitchell's Plain (KMP) Survey respondents refused job offers, even with severe levels of unemployment.



Source: Encyclopaedia Britannica Inc., 2013

South Africa offers a unique opportunity to investigate the effects of commute costs on unemployment. Transport costs can be measured in a number of ways. For instance, people need resources to search for jobs, and one of these resources is the money required to commute. Transport costs are a vital part of job searches and commuting to work. They are even more crucial for unskilled job seekers living on the edge of urban society where they may not have access to the Internet, labour market networks and information linked to potential jobs. Patacchini and Zenou's (2006) paper develops a simple model for optimal search intensity by using sub-regional panel data from England from 1995 to 2000. They find that decreased search costs and the future reward of possible employment increase the intensity with which the unemployed search for jobs. They argue that the need to search for a job increases as the cost of living rises. This in turn makes it relatively less costly to search for a job today. As the cost of consumables becomes a greater burden, so does the need to search for a job. South Africa's employment landscape is very different to that of England, and unemployed individuals are likely to incur far greater costs in trying to overcome larger distances to search for jobs. The benefits of decreased search costs, similar to Patachini and Zenou's, are also likely to be far greater than for an individual in England who has to cover less distance. However, the large non-economically active (NEA) portion of the South African population suggests there are other factors at play. Furthermore, why are there so many discouraged workers? Do they not feel the pressure of securing a job? Measuring the relationship between transport costs and job searches could reveal that increased expenditure reduces incentives for job searches. Rising transport costs can also inflate reservation wages. This is because individuals raise their reservation wages in a bid to recoup some of the search costs they have incurred.

A small number of studies have considered the relationship between transport costs and labour market outcomes in South Africa. The literature on the types and scope of effects that monetary commute costs have on unemployment is scarce. Studies have investigated related issues, but none have used reported monetary costs. Instead, many studies have used a proxy for commute costs. This is because transport cost measures are not very good or they are limited by data. For instance, Hinks (2008) uses distance from public transport to argue that better access to transport facilities may increase the

potential searching radii of job seekers. The lack of transport cost studies does not necessarily reflect the unimportance of transport costs, but rather the lack of access to transport cost data. Hinks (2008) uses the Labour Force Survey (LFS) as a data source and notes that his model omitted public transport costs because they were not included in the data set. Kingdon and Knight (2006) also use proxies for transport costs. Their indicators of remoteness measure the impact of spatial deterrents on job searches.

This paper improves on these past works by using better measures of transport costs. These improvements are conducted in two main areas: the transport and job search relationship, and the transport cost and wage relationship. The first analysis builds on the work of Kingdon and Knight (2006) and Hinks (2008). This paper is innovative because it uses a more comprehensive measure of transport costs and a newly released data set. Indicators of remoteness, similar to those used in Kingdon and Knight (2006), are also included as proxies for job search costs. The second part of the analysis explores transport costs and their effect on real wages (i.e., net transport costs) and reservation wages. The second approach is, to my knowledge, entirely new in the South African literature.

This research uses a similar methodology to that used by Kingdon and Knight (2006), but uses the 2009 National Income Dynamics Study (NIDS) and Labour Market Entry Survey (LMES) data sets. This paper is innovative because it uses a more comprehensive measure of transport costs and a recent data set. This paper uses indicators of remoteness, similar to those used by Kingdon and Knight (2006), as proxies for job search costs, as well as several other new indicators. The research report is divided into six chapters. Chapter 1 introduces the research and reviews the relevant literature. Chapter 2 reviews relevant analyses and previous work relating to unemployment in South Africa and the determinants of job searches and transport costs; and provides an overview of the data sets used. Chapter 3 presents an analysis of transport costs and their potential effect on unemployment through job searching. Chapter 4 examines the relationship between transport costs and reservation wage determination. Chapter 5 explores the relationship between transport cost premiums and transport

costs, while focusing on time as a proxy for transport costs. The paper ends with some concluding remarks.

1 Chapter 1

1.1 Literature review

Search theory and its development continues to be motivated by the simple notion that workers generally have access to a variety of job opportunities and seek the best one given many alternatives (Mortensen, 1986). Understanding why millions of people are unemployed is important because doing so allows labour economists and policy makers to make informed decisions about these under-utilised resources. In this research, this is done in a framework using the search theory of unemployment. This framework is a large departure from the classical labour market theory of employment. The notion of a central market where employers and workers gather to trade labour at a single price is too far removed from reality to have any bearing on search theory. Furthermore, classical theory imposes counterintuitive assumptions, like assuming that firms and workers have complete information on job availability, which ultimately means that there is no need to search for work since the stochastic nature of the process is removed (Fitzgerald, 1998). Classical theory also requires that the demand and supply of labour are at equilibrium and deviations from this equilibrium results in what is termed involuntary unemployment (Fitzgerald, 1998). Hence, the classical notion of labour market interaction is considered frictionless. In contrast, the search theory of unemployment accounts for the uncertainty associated with searching and acknowledges that the process costs time and money, and attempts to identify why firms have vacancies while there are millions of unemployed job seekers. This is done by explicitly modelling these trade frictions. In a context like South Africa, where there are likely to be large frictions within the labour market, it is essential to take these into account.

The introduction of information costs into search theory was pioneered by Stigler (1961). He recognised that information is imperfect, hard to obtain and costly to both competing firms and job seekers. In Stigler's subsequent work he argues that as young workers enter the labour market, it may seem as though there are little to no vacancies for employment. However, in reality, millions of unskilled and semi-skilled jobs are available in almost any economy and thousands of jobs are available to highly skilled individuals (Stigler, 1962). Although this may not be entirely true given the recent hardships that have befallen many economies, the lesson remains true—ignorance is costly. As

Stigler (1961) aptly points out, the dominant cost associated with searching, from the perspective of both the employer and potential employee, is time. The same is true for firms. For the job seeker, this cost can be approximated to the number of potential employers canvassed. Furthermore, the optimal search occurs when this search cost equals the expected marginal return of the searcher (Stigler, 1961).

Stigler (1962) aims to determine wage rates. He uses a sample of wage offers given to a small group of accounting graduates from a single university to illustrate price dispersion of wage rates under the assumption of homogeneous goods (the graduates). Given perfect information among buyers and sellers, firms and graduates respectively in this example, one price exists for a given commodity (Stigler, 1962). Perfect information, however, exists only in the classical framework and is rarely, if ever, possessed by anyone or anything. This is because an individual's marginal costs exceed their marginal benefits when trying to gather information on alternative prices (Stigler, 1962). Even with this biased sample, Stigler finds a 5% to 10% dispersion in hiring rates among this group. Search theory centres on marginal benefit analysis because workers search until they find wage offers where the expected marginal benefit equals the incurred marginal cost. Stigler (1962) asserts that the greatest gains from searching are achieved in earlier periods because current searches informs future searches. However, this is only true when the possibility of successive wage offers, based on searches, exists (Stigler, 1962).

One of the main themes of this paper, discussed more thoroughly in Chapter 3, is whether transport costs affect job searches. Research as early as Stigler (1962), has found that at various wage offers made by employers, fewer searches are undertaken if the search costs are large. Attributing these costs to transport costs alone is challenging. Searches conducted by the unemployed or workers are less costly when employers are easily identified because it reduces the possibility of fruitless inquiries (Stigler, 1962). This means that the more easily workers identify employers, the smaller the wage difference (dispersion) of a sample. Given this knowledge, it is possible to construct an indicator that measures the ease with which workers can find employment by simply calculating wage dispersion (standard deviation of wages) while controlling for skill. For both searchers and employers, gathering

information about each other amounts to a capital investment (Stigler, 1962). For employers, gains from information are acquired by attaining the most suitable worker (in terms of productivity) for a given wage rate. Increased search by the unemployed or workers yields better information about a single job at multiple wage rates displayed by employers. This information then becomes pervasive in the economy. Therefore, for a given economy, increased search intensity benefits all who enter the labour market (Stigler, 1962). However, costs increase with search, so the returns to searching display diminishing returns. Ultimately, the individual faces a sampling size problem in terms of trying to determine which employers to access (Mortensen, 1986). The individual chooses an employer based solely on the largest random wage offered in the sample. In South Africa, discouragement in the labour market is large. This suggests that, for those who seek employment, their gains are far less profitable because less information about jobs is available. This could create a spiral of discouragement and exacerbate the unemployment problem even further. If the value of information is less than the cost of attaining it, then individuals reduce their search activities and resort to passive searching.

Job search theories that followed Stigler's approach view the worker's decision as either continuing to search or to stop. This was captured by sampling wage offers successively from a randomly obtained sample of employers. Stigler's (1962) model is static and takes place in a single instance of time. This is because it is non-sequential in nature, which means that the numbers of offers are determined before the search begins. A sequential strategy revolves around the decision to accept an offer or reject it and continue searching, and it is generally considered superior to its counterpart. McCall (1965) was one of the first to consider sequential searching. Imagine an individual who enters an area seeking employers, and when presented with an offer decides whether to stop or continue searching. The sample size of employers is limited, in part, by the individual's stopping rules. The individual's wage expectations are primarily governed by their skill level and experience, and any wage offer that fails to meet these expectations falls away and the search continues. The more costly information regarding wage offers is, the more likely the individual is to decrease their expectations and search activities. In terms of this model, regardless of the environment, the job offers that the searcher

obtains arrive at random intervals. In reality, the random nature of the sample size can be interpreted as a search spell with a random duration (Mortensen, 1986). These search theoretic analyses are generally considered more realistic with a higher expected return to future incomes and run counter to Stigler's (1962) work, with a static employer sample size.

No rational worker (or work seeker) would ever wait indefinitely for the best employment opportunity because search costs rise with this indefinite wait and because of the general psychological phenomenon to value immediate returns or gains more than future ones (Mortensen, 1986). This time discounting helps to explain the way individuals perceive the results of their efforts in the present and future. It assists in understanding why many individuals choose to procrastinate, even if the future benefits of, for instance, creating a daily or weekly work schedule would make the individuals better off. This is because the perceived benefits of these tasks or decisions are not immediately apparent. The development of search theory is a historical weave of contributions made by various authors from as early as 1834, with Scottish economist John Rae who examined this inter-temporal choice in terms of conflicting psychological motives. He augments inter-temporal choice models by including psychological motives to understand capital allocation and the determinants of its production. Samuelson (1937) later summarises all these psychological motives by using a single parameter, known as the discount rate, in his discount-utility model (Frederick et al., 2002). At the time, the model was thought to be an accurate theoretical description of individual behaviour, and the discount rate summarised all the different choices that could be made. However, Samuelson did not share this view and many anomalies and inconsistencies were later found with the model, even though it is still used in economic research. One of the major failings of the model is that it does not consider how an individual's personal understanding of their own preferences could affect their inter-temporal behaviour (Frederick et al., 2002). From 1999, studies began including individual specific knowledge about preferences into this model because of the failure of empirical research to establish any stable estimates of the discount rate (Frederick et al., 2002).

Since the late 1970s job loss and voluntary quitting were included in the search theory framework. Lucas and Prescott (1974) made the first complete contribution to this topic, which served as a basis

from which many authors extended their analyses, most of which were completed by the early 1980s. For instance, Mortensen (1986) extends the Lucas and Prescott model to find the equilibrium employment level, including search unemployment and non-participation in yet another variation of the original standard search model. By using fundamentals of the Markov chain model, an individual's historical transitions into different states (non-participation, employment and unemployment) and the standard search model can be used to explore differences in participation and unemployment rates. Models within the search theory framework that analyse the transition from one of two states are considered two-state models (Mortensen, 1986). For instance, in terms of job turnover, the focus is on the individual's probability of transitioning from one job to another. Furthermore, the transition from unemployment to employment in the original search model hinges on the individual's stopping rules that govern their decision to accept or wait for a job. Mortensen (1986) asserts that this analysis can be expanded to create a three-state model, but this model characterises an individual's labour market experience and the two-state version.

The impact of age and job tenure are extremely important in studies conducted before the 1980s. These ideas still persist in the recent literature. Both have been found to have a negative relationship with job turnover. Quitting often occurs during the early stages of employment when a worker might realise that the characteristics of the firm are not well suited to them. However, once permanent employment ensues, it is hard to justify the exit of an employee who has remained with a firm for some time. During this time, the individual has experienced firm-specific work and training that make them better than any other alternative (Burdett, 1978). Furthermore, wages and age increase with tenure, since the individual's productivity rises over the same duration, and increased wages and tenure require increased productivity.

Search theory centres on the premise of understanding the wage dispersion that exists between homogenous workers and a framework explicitly designed to capture the differences caused by trading frictions that are generated from a trading process between two parties. In the basic example of an individual searching for work at a fixed rate over time, a wage offer that is greater than their reservation wage is all that is required for an optimal decision. Given that the wage offer sampling

takes place at a fixed rate, the example implies a random length of unemployment (or search duration), number of firms, and wage offer (Rogerson and Wright, 2005). This basic model is only meaningful, in terms of motivation for search theory, if there is wage dispersion and if the dispersion exists in the presence of homogenous workers. The latter point is critical to search theory because wage dispersion can easily be generated for workers with different productivities. Hence, wage dispersion is only important for a given *type* of worker (Rogerson and Wright, 2005).

A number of important distinctions need to be made with regard to sequential models and time horizons before discussing the specific extensions of the standard search model. Sequential models are generally dealt with either within an infinite or finite time horizon framework. With an infinite time horizon, in a very simple example of a basic search model, when an individual seeks employment he ventures out and each period (day, week, etc.) he generates a single job offer. In a model with no turnover the job seeker remains in this job forever. Furthermore, an infinite time horizon assumes a stationary reservation wage, which generally means that if an offer does not meet the particular stopping rule it never can and will be refused indefinitely. Within the finite time horizon framework, a job seeker determines a maximum number of possible jobs which he could possibly search and searches in a manner so that all possible offers unfold sequentially. This process only stops once the stopping rule is met and yields the maximum expected gain. Further distinction can be made within the two time horizons with regard to job offer recollection. The difference between infinite and finite time horizons in this regard is that job offer recall is generally dealt with in the latter framework. This is because recalling the highest past job offer (current or last offer) makes no difference to the outcome of search, in the infinite time framework, which continues until the critical value (the reservation wage) is exceeded (Lippman and McCall, 1976).

1.1.1 Exogenous wage dispersion

The seasonal or time series behaviour of unemployment rates is well documented from various data sources, showing that unemployment levels are cyclical in nature and closely related to the business cycle. Furthermore, differences among demographic groups, in terms of unemployment and participation, reveal that a worker's labour market history can be viewed as a random process

(Mortensen, 1986). According to Mortensen, this idea prompted significant developments in search theory and its analysis of individual labour market experiences, which further contributed to the search theoretic approach of analysing unemployment spell durations. The original wage search model was created from sequential statistical decision theory (Mortensen, 1986). Mortensen (1986) expands on the original form by designing a model that focuses on employment searching in decentralised labour markets and acknowledges that information signals—with respect to job vacancies and compensation—are, at best, imperfect. He extends the standard wage search model to endogenously include effort levels for on-the-job searches. By doing so, he explains what is commonly seen as the natural tendency for individuals to switch jobs when their wages decrease. Generally, higher paying jobs have better incentives for individuals to remain entrenched in their jobs because of the benefits associated with increased tenure, which is usually not the case with low paying jobs.

Regarding the evolution of the search theoretic approaches to analysing unemployment, in terms of spell duration and the rate at which employers hire and fire workers, most economic studies leading up to the 1980s are modelled on an individual's employment or unemployment history using a Markov chain (Mortensen, 1986). This is largely based on longitudinal observations which show that, historically, as individuals in the labour market move from one state to another (e.g., from employed to unemployed) they only consider their current situation or state. Therefore, the outcome of their current state affects the outcome of their future state. Given infinite state possibilities, the process of shifting into different states is considered a random and memory-less pattern. However, changes to this modelling approach began when an individual's response to random changes in their labour market opportunities and opportunity costs were used as the basis for determining their movements across enumerable states (Mortensen, 1986). Mortensen (1986) explains that reservation wages should decline with search duration because of two phenomena. The first is that the probability of finding a job could diminish with time, given that missed job offers are likely taken up by other members of the labour market and the ever-mounting psychological effect of fruitless searching. The second is the constrained access to money required to fund such searching—these individuals are likely to acquire

monetary assistance from their households and possibly their communities. Their likelihood of attaining finance from a bank is slim, given their high probability of defaulting. These over-arching problems may force a job seeker to either accept any job that compensates for the value of leisure lost or resign into the ranks of the discouraged. Mortensen's review of the literature is confined to the decision theory framework of sequential searches.

Early contributions to equating reservation wages among the employed and unemployed began with Burdett (1978), who explores the differences between both groups in terms of search costs. Labour economics and job search studies that precede Burdett's work assume that once a worker has accepted a job offer at least equal to their reservation wage, they remain in service until retirement or death, because the costs of searching for better employment are too high. However, Burdett shows that this only holds if the cost of searching is greater when employed. Under these conditions, an employed individual who is looking for another job will migrate towards any job that offers wages greater than their current one. Mortensen's on-the-job search model formulation assumes away any difference that may exist between being jobless or not, and with this asserts that job acceptance is determined by the job that compensates for the lost leisure time value. Once accepted, most individuals pursue higher paying jobs (Mortensen, 1986). Applying a binary approach to search behaviour makes it easier to decipher the sequential search model. In reality, search intensity is an important component that determines unemployment duration. Efforts to find employment are increased primarily to reduce the length of the unemployment spell by increasing the probability of finding an acceptable or better wage offer (the latter applies to those already working). However, this relationship is not linearly positive because longer durations of increased search efforts are likely to garner increased search costs, in turn producing diminishing returns (Mortensen, 1986). Mortensen (1986) includes search intensity in his search-on-the-job model by assuming that wage offer arrival rates are proportional to increased searching, while simultaneously making search costs an increasing convex function of search efforts. In doing so, he identifies the optimal search effort that equates its marginal cost and return. Hence, an individual seeks employment when their marginal benefit, measured in leisure value terms, is greater than their marginal search effort cost. They will continue to do so until a sufficiently well-paying job

reduces their optimal search effort to almost zero. This marks another deviation from the standard search model because it provides an alternative explanation for the positive relationship that exists between wages and years of work experience (Mortensen, 1986). Prior to Burdett's (1978) paper, the dominant assumption in the quit rate literature was that increased age and job tenure decreased the likelihood of quitting. This is because it was thought that productivity and work experience was a consequence of accumulated firm specific capital, that is, working and training (Burdett, 1978). The alternative view states that rising wages are due to longer experience and the resulting increased probability of finding a high paying job (Mortensen, 1986).

The standard search model continued to develop by assuming that the job seeker has full information about the characteristics of the job. In all fairness, some surface level information regarding the job must be known, otherwise there would no way to differentiate one job from another. This is primarily conveyed through the wage offer, but even within job categories, small differences exist between subcategories that a single skill can perform fairly well at. However, many characteristics of a job, particularly future income streams, can only be gained through intimate knowledge and personal experience. The decision to stay or move is constantly re-evaluated as more time is spent on the job, and a move results when an individual decides that a job is not conducive to their future wellbeing (Mortensen, 1986).

McCall (1970) employs this analysis and uses a simple search model to show that the optimal strategy for a searcher is to accept or reject a wage offer at some critical value—the reservation wage (strategies that incorporate this exhibit the reservation wage property). He analyses a simple search model to distinguish between the discouraged and frictionally unemployed. His model shows that reducing search costs decreases the number of discouraged, which is similarly achieved by increasing the skills of the individual. This is in line with Burdett's (1978) analogy of firm specific capital accumulation and increased wages through tenure. Furthermore, a more general version of this model is created to incorporate employment duration into the cost benefit problem that searchers face. This is achieved by including a discount rate over multiple time periods while maintaining the

fundamentals of a Markov chain. Both models produce the same optimal search strategy results for either group.

1.1.2 Endogenous wage dispersion

Much of the job search literature in the 1970s and leading up to the 1980s was criticised for its inability to explain wage dispersion in markets that called for endogenous firm behaviour. Diamond's (1971) paper is considered the most dominant critique of job search theory. Diamond showed that while endogenous firm behaviour persists, no wage dispersion exists in equilibrium because no firm is willing to offer wages greater than the reservation wage (Postel and Robin, 2002). At the same time, any firm at equilibrium offering wages less than the reservation wage will have no employees (Burdett and Mortensen, 1998). The resulting wage offer distribution is narrowed down to a single monopsony wage, which ultimately means that all analyses in the field of job search theory must generate equilibrium wage dispersion to avoid being relegated to a simple labour market monopsony model (Postel and Robin, 2002). Diamond's (1971) conclusion of the basic search model is severe and questions the legitimacy of search theory because he disproves that, even with heterogeneous firms, search frictions do not yield a non-degenerate wage distribution, that is, a unique wage distribution solution. These criticisms had a major impact on analytical work that used standard (static) search theoretic frameworks during the 1970s. Kiefer and Neuman (1979) suggest that this slowdown, and lack of empirical work, was due mainly to the difficulty associated with introducing the sequential decision process. The basic search model is meaningful only if it is used to identify the market conditions required for homogeneous workers to generate this unique solution. Wage dispersion motivates search in the first place. This meant that any economist who wanted to continue using the search theoretic framework among workers needed to generate a non-degenerate wage distribution by finding some wage dispersion in equilibrium. This was undertaken by many economists and demonstrated in a number of ways, some of which have been discussed. Albrecht and Axell (1984) and Burdett and Mortensen (1998) both introduce endogenous wage dispersion into their models by allowing for worker heterogeneity and on-the-job search, respectively.

Albrecht and Axell (1984) develop a simple equilibrium model of search unemployment based on the standard search theoretic framework, namely, the rational decision of an individual to accept a wage offer or continue searching for more appealing offers. This model generally finds that increases in unemployment compensation should lead to a net reduction in search costs, increasing both reservation wages and search duration. This ultimately leads to an increase in equilibrium unemployment. Two problems with the standard framework motivated their work. First, there is neither evidence nor a general equilibrium model that justifies a unique wage offer solution from sequential searching. Second, the standard model requires an exogenous and unchanging wage offer for a static comparative analysis of changes in unemployment compensation. This is contradictory because changes in the latter would change the wage distribution, which should, in theory, be exogenously determined. They focus on the effect of unemployment compensation and its effect on equilibrium unemployment by using a model with wage offer distributions that are endogenously determined. This is achieved by including heterogeneous workers. Albrecht and Axell (1984) construct their model by first defining individuals and firms that maximise utility and profit, respectively. There are two classes of individuals, those with high and low tastes for leisure, which are subdivided by their employment status (employed or unemployed). Here, reservations wages equal the value of continued searching, which is determined by the random draw from the two-point wage offer distribution, the dichotomous leisure values and the risk of death. They describe two types of firms: active and inactive. Active firms service searching individuals with wage offers. They then define the equilibrium unemployment rate based on the number of individuals entering the economy per period, the ratio of individuals accessing active firms who are limited by the per-period labour supply, their own productivity and wage offered. With these tools, they establish conditions for a two-stage wage equilibrium, which they use to examine the effects of various unemployment compensation levels on the equilibrium wage distribution. They find, in general, that an increase in unemployment compensation increases the equilibrium unemployment rate. This is in line with the standard search theoretic framework, with one exception. For individuals with low leisure values, the opposite is true, that is, increases in unemployment compensation decrease the equilibrium unemployment rate.

Following from Albrecht and Axell's (1984) introduction of wage dispersion through worker heterogeneity, this section discusses a paper considered by many to be the holy grail of search theory as well as a solution to Diamond's (1971) critique. Burdett and Mortensen (1998) show that wage dispersion can exist at the equilibrium, even with homogenous workers. They assert that all studies preceding their paper do not explain the stable wage differences that persist among heterogeneous workers across industries and employers. They show that wage differentials persist in markets where rational utility maximising individuals constantly seek better wages by finding the next best job. This is true whether workers are homogenous or not.

Burdett and Mortensen (1998) combine labour market models and standard search theory by treating search friction as a form of information regarding labour terms and job availability. They do this by slowly building a model from one with simple implications and applications to various theoretical and empirical predictions, with respect to employer and worker behaviour in the labour market, to a more generalisable model. To start, a simple model is created to analyse wage dispersion within an economy where both firms and workers are homogeneous. In this version, Burdett and Mortensen focus on steady state behaviour within the standard search theoretic framework, where individuals randomly sample wage offers in a sequential fashion. They calculate the steady state number of unemployed workers, which is used to calculate the steady state number of workers available to a firm offering a competitive wage. These wages can be counter offered in a wage posting game, in which alternative offers allow employed workers to move to ever increasing wage employment. They use the profit maximising solution of the firm, the reservation solution and non-degenerate wage distribution to find a unique equilibrium. Burdett and Mortensen (1998) even determine Diamond's (1971) monopsony wage solution since their model is characterised by a positive relationship between staff retention and higher wage offers. This implies that larger firms experience lower quit rates (Burdett and Mortensen, 1998). The model further implies that higher wage earners are likely to be those with longer tenure, and more experience, because frictional unemployment only exists when workers respond to higher wage offers from employers. They then generalise the model by introducing worker and employer heterogeneity successively in two separate steps. By allowing

workers to value their leisure differently, Burdett and Mortensen include the differentiated opportunity costs of employment among individuals as a determinant of wage dispersion. In such a case, some individuals who are efficient at their jobs will not find matches. Thus, they find an equilibrium solution to the wage posting game. By introducing firms with differentiated productivity levels, Burdett and Mortensen drive the point that higher wage offers increase the workforce size, reducing the number of quits in large firms. Both predictions can be applied to general observations of firm and individual behaviours.

Given the high levels of discouragement and unemployment and the geographical distribution of unemployment in South Africa, it is likely that the search behaviour of individuals follows a sequential pattern. It is unlikely that the average unemployed individual has enough knowledge of employment opportunities (or employers) to pursue a static search, as described by Stigler. Given the limited information regarding employers, employers are basically randomly distributed and therefore job offers are likely to arrive at random intervals. For those willing to search, the undertaking is costly and success is largely based on factors that contribute to their minimum reservation wage criteria. These include monetary costs, skill levels and a range of psychological factors, making these individual heterogeneous. The costly nature of sequential searching and the lack of information regarding employers invariably mean that, for many, the costs of searching outweigh its benefits.

1.1.3 Models with constant wages (degenerate wage distribution)

The equilibrium or natural rate of unemployment is the theoretical case that exists when there is a difference between those who are employed and those who want to be employed at some prevailing market-clearing real wage. Formally, these two groups are known as the labour supply and the labour force. Under these conditions, equilibrium employment exists at the point where labour supply and demand intersect and involuntary, or forced, unemployment does not exist. The natural unemployment that persists, exists only as frictional or structural unemployment. Hall (1979) uses the theory of employment duration to find the equilibrium unemployment rate, where jobs and workers are perfectly homogenous. This implies a degenerate wage distribution because there is no wage variability. The theory assumes that unemployment spells are short lived and employment is the result

of efficient negotiation between employers and employees about the duration of their term. Hall shows that the natural rate of unemployment is independent of labour supply and demand, and that the costs incurred by employers and job seekers are the only real determinants that strongly affect, what he terms, the efficient duration. Unemployment duration, Hall asserts, is identified through job separations in the form of temporary layoffs and labour market entry. Furthermore, the main reason unemployment exists is the prevalence of short term contracts or, put differently, the lack of permanence in the work place.

Hall adapts the standard employment equilibrium model by augmenting it with his own efficient separation rate and efficient job finding rate, both of which are identified using isocost and indifference curve analyses. He constructs new labour supply and demand curves using these new variables and finds that the socially optimal level of unemployment is quite low and that high recruiting costs are the main reason that jobs are not abundant.

Pissarides (1979) develops a search model that includes a centralised employment agency and random searching. An individual must choose between these two search methods. He uses this model to derive full market equilibrium for both the firms and the labour market. He argues that there is more to unemployment and searching than just the reservation wage, which was the main focus of search theory when his work was published. The decision to reduce or extend the search duration is governed by the intensity of search methods used by the individual, that is, random searching or registration with an agency (Pissarides, 1979). To rely less on the reservation wage in his analysis, Pissarides simplifies his model by assuming that no wage variability exists, that is, a degenerate wage distribution. This means that the analysis focuses on vacancies, rather than the highest paying jobs that meet individual minimum criteria.

By using a degenerate wage distribution together with other endogenous parameters, including unemployment, random searches, and registered and unregistered vacancies at agencies, Pissarides (1979) analyses the steady state equilibrium. He finds that reducing the cost of random searches

decreases steady state unemployment while increasing job matching. Ambiguous results are found for cost reductions at the agency level (Pissarides, 1979).

1.2 Summary

This chapter attempts to understand the fundamental factors that contribute to search theory and finds that at its core, search is governed by information which is imperfect, costly and hard to obtain (Stigler, 1976). How this is modelled has changed over the years, from static to sequential models, but generally studies still centre on the marginal benefit analysis that is intrinsic to the search strategy of the individual and employer. Equating the benefits of search to its cost indicates that individuals have some sort of stopping rule where the costs outweigh the benefits of continued search. All else equal increased search intensity benefits all who enter the labour market because information becomes pervasive. However, there are diminishing returns to search given that costs increase with continued search. Attempts to measure this requires understanding the wage dispersion among workers, heterogeneous or homogeneous, which is created from the frictions that are generated from the trading process between two parties.

2 Chapter 2

2.1 Unemployment in South Africa

The South African economic landscape is characterised by rural-urban migration (Leibbrandt et al., 2002) and, what the Department of Environmental Affairs calls, low density urban sprawl. This urban sprawl separates job seekers from their potential employers as they move further away from urban areas. Leibbrandt et al. (2002) suggest that this urban sprawl may occur because many of the young, rural unemployed survive by staying close to pensioner households that are often located on the urban fringe or in rural areas. This is largely due to the spatial consequences of apartheid that intentionally forced none white people to the periphery (urban fringes and rural areas). Living on the urban fringes may increase the search costs (including transport costs) of attaining work resulting from the increased commute length to and from urban labour markets (Leibbrandt et al., 2002). Coupled with increased search and transport costs, this makes migration to urban areas a risky venture for job seekers (Leibbrandt et al., 2002) and inevitably increases urban sprawl. Leibbrandt et al. (2002) also find that rural migration by these participants is less common since they do not have a decent chance of attaining urban employment. The premium prices paid for urban property results in poorer job seekers remaining on the urban fringe and bearing higher transport and search costs.

This legacy has left rural communities in a difficult situation regarding the decision to find work in urban areas because they must either physically move to those urban areas or endure a daily commute. Seekings et al. (2004) suggest that rural communities with a history of social networks, education and entrepreneurship may have created migration linkages between these rural communities and urban areas, making them less vulnerable to poverty after apartheid. Nattrass (2000b) examines interviews from unemployed people who stated that they remained in rural areas because of the lack of social contacts within bigger cities and towns. Nattrass (2000b) also finds that rural urban migration tended to occur among people who had contacts in urban areas.

Kingdon and Knight (1999) find that for regions in some OECD countries a negative relationship exists between unemployment and wages. Their paper tests the veracity of this relationship given the

high levels of unemployment in South Africa and is similar to Blanchflower and Oswald's (1994) work. This is in contrast to the conventional (or neoclassical) supply and demand model of the labour market that generally asserts that higher wages facilitate higher unemployment (Kingdon and Knight, 1990). In the neoclassical model, unemployment occurs as a result of labour supply exceeding labour demand. This is because firms are unwilling to hire at wages above the market clearing level, which drives down employment (Blanchflower and Oswald, 1994). Kingdon and Knight (1999) assert that a positive relationship exists between searching and wages as job seekers try to recoup high search and related costs. This has been described as the wage curve or the conflict wage curve (Blanchflower and Oswald, 1994; Palley, 2007). In effect, the real wages are indirectly determined by the wage bargaining power of employees (or potential employees) and not by their marginal product of labour (Palley, 2007). According to this argument, if Region A has high levels of unemployment and Region B has lower levels of unemployment, a worker in Region A would earn lower real wages than an equivalent worker in Region B. This is because in Region B, workers have greater bargaining power, because of the region's higher job availability. A region with lower levels of unemployment has more available jobs. With this knowledge, when job seekers who are looking for a particular job or wage are confronted with an employer who is unwilling to pay that wage, they will move on to find someone who will. In effect, unemployment behaves as a disciplining mechanism, making it unnecessary for firms to pay high efficiency wages (Kingdon and Knight, 1999). One justification, made by Blanchflower and Oswald (1994), is the involvement of trade unions during periods of high unemployment and their increased concern for employed and unemployed members. They state that it is possible that unions may be willing to negotiate lower wages in the face of high unemployment and job losses. In the South African context, wage rigidity is partly due to unwavering trade unions Blanchflower and Oswald (1994).

Given the South African scenario of high unemployment and particularly high unemployment in rural areas, job seekers, in a bid to try and recoup their search and transport expenses, require (signal) higher wages and show an unwillingness to take up employment unless these wages are met. In this instance, the wage curve does not exist. However, it is unclear whether this contradiction only occurs

if employers pay that wage or if the mere signalling by job seekers is enough. This depends on whether institutional bodies, such as wage bargaining councils, exist and whether they can exert enough force on the minimum wage. If employers cannot find employees willing to accept a low wage, they must settle for these higher wage job seekers. South Africa is well known for its trade unions and their role in driving up wages. Kingdon and Knight (1999) state that wage boards and bargaining councils play a large role in wage determination. They assert that African unionised workers receive wages that are 19% higher than those of their non-union counterparts. However, they also claim that this distinction is less pronounced in areas where employers are affiliated with Industrial Councils (ICs). ICs consist of a collection of employers in a sector who agree to negotiate on numerous issues, including the minimum wage. Minimum wages for a sector are established through collective bargaining. Once an IC is formed in a particular sector, their prescriptions become enforceable to all firms, or employers, within that sector and heavy penalties are handed out to those who disregard them (Kingdon and Knight, 1999).

Kingdon and Knights's (2006) paper focuses on whether discouraged people are unemployed (broad definition) or out of the labour force. Ultimately, their paper aims to ascertain whether the discouraged are treated as part of the labour force, thus shedding light on the disparity between the definitions of broad and narrow unemployment. They achieve this by formulating three new tests centred on the questions of whether non-searching people are richer and happier, and whether they have a smaller impact on local wages than searchers. Kingdon and Knight (2006) assert that South Africa has a notably high unemployment rate, but is also a country where discouraged people are prevalent. They note that discouraged people are no different, in terms of social loss, than the searching unemployed. This is because people who want and are willing and able to work would seek employment if employment was abundant (Finegan, 1978). For this and other reasons, this paper uses the broader definition of unemployment.

Comparing discouraged and searching workers in terms of their transition into employment may be more difficult under conditions of high employment (Kingdon and Knight, 2006). This may be due to the notion that the discouraged may see relative gains from not searching, since the costs to the

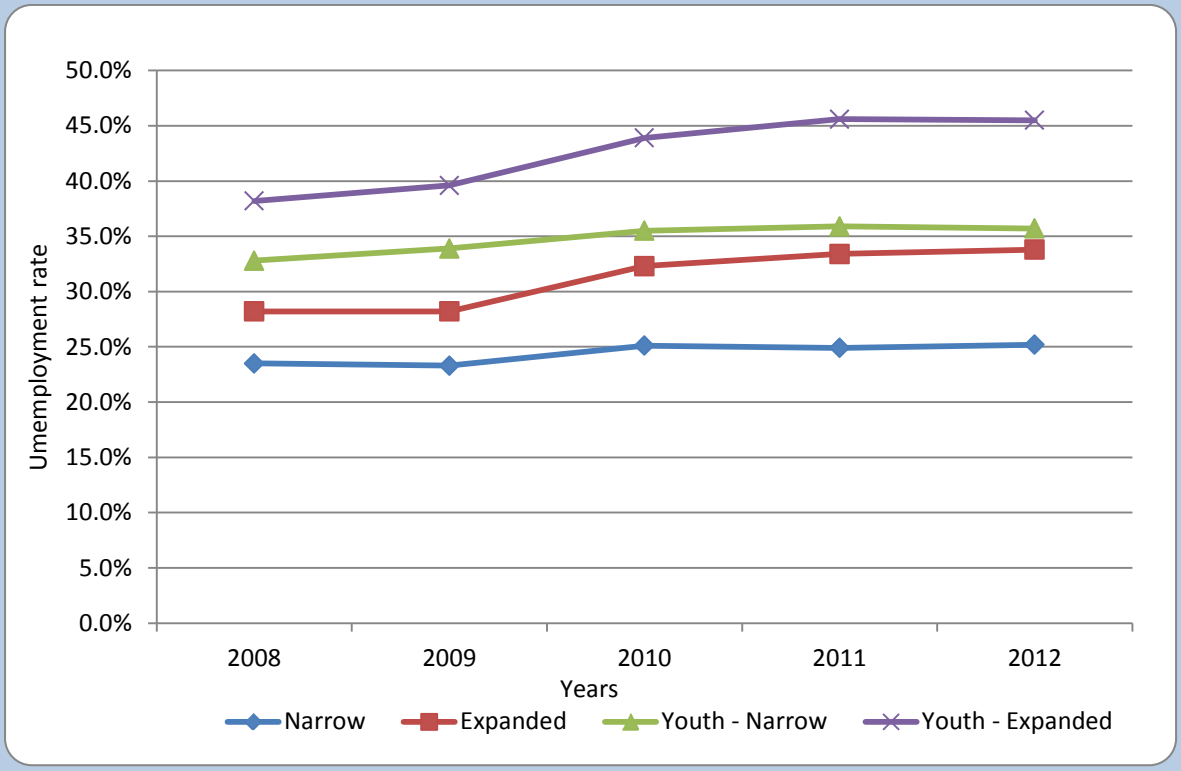
searcher are much higher under these conditions. Kingdon and Knight (2006) make the reasonable assumption that searching, or not searching, is based on an individual's taste for leisure and the small gains they think they attain from searching. These gains are even smaller given the prevalence of unemployment in South Africa. If an individual relies on passive searching, they will learn of fewer jobs and indirectly (or inadvertently) decrease their direct search costs as well as the probability of attaining a job. Passive job searching is generally chosen if the individual incurs negligible costs while acquiring some future benefit in terms of job vacancies. Falling passive search costs (direct costs) raise the reservation wage. This is because the individual believes that they can continue to search until they get the job that they desire. Passive job seekers will only move to active search behaviour if the benefits are greater than the cost incurred or if they learn about other possible vacancies through increased job offers. Reservation wages are typically smaller for active unemployed job seekers, compared to their passive counterparts, because these people are more willing to accept lower paying jobs to recuperate losses created through searching. Those that are discouraged need not recuperate anything, since their passive search costs remain the same. Hence, the higher the reservation wage, the better the benefit to cost ratio of passive search for the unemployed individual (Kingdon and Knight, 2006).

Kingdon and Knight (2006) provide two interpretations of the reasons why discouraged individuals are in this discouraged state. First, individuals choose unemployment because the household is the provider and supports poorer members. They refer to this as the taste for unemployment hypothesis. Given this support, needy household members have an incentive to avoid searching, and higher household incomes perpetuate this behaviour. Second, the discouraged worker hypothesis describes individuals who are reluctant to search because of high unemployment rates and hence, have a low probability of attaining work. Ultimately, this creates a low cost to benefit ratio, which is much lower given the prevailing high unemployment rate. Kingdon and Knight (2006) find that the unemployment gap, which basically describes the discouraged, and duration clearly illuminate the depth and breadth of unemployment in South Africa. Using the October Household Survey from 1994 to 1999, they show an average unemployment gap of around 13%. This accounts for about half, more in some

years, of the official (narrow) South African unemployment rate, if the broad rate was used instead. Their results also show that 67% of these individuals have been unemployed for at least a year, with 37% of the individuals having been unemployed for three years, before getting jobs.

Kingdon and Knight (2006) evaluate their research question in terms of happiness by primarily using an ordered probit model. They create an index of happiness using five rising states of happiness as the dependent variables and find that unemployed non-searchers and searchers are equally unhappy. This study was conducted in South Africa as well as other countries. They also find evidence to warrant including the non-searching into the definition of unemployment. This is because their results suggest that an increase in the broad unemployment rate increases the narrow to broad unemployment gap, causing the unemployed to become discouraged workers.

Figure 2.1 highlights these issues by showing unemployment in South Africa in the first quarter of each year from 2008 to 2012. There was a sharp rise for all unemployment types from the end of 2008 to 2010. This was largely due to the recession during those years. These effects are still being felt today, with small declines in employment in 2012.



Source: Development Policy Research Unit (2012)

Figure 2-1: Unemployment in South Africa from 2008 to 2012

Within this milieu and the context of the growing concern over the inability of the general population to achieve employment status, the following chapters explore a key potential barrier to employment in South Africa, namely, transport costs. The paper examines whether high transport costs limit take-home pay and whether this reduces the incentive for people to accept jobs. Unemployment is a phenomenon that changes over time, persists worldwide and differs from area to area. Even in regions that are fairly close to each other, there will be marked differences in multiple facets of economic life. To a large degree, unemployment is governed by national, and not local, decisions (Aragon et al., 2003). Aragon et al. (2003) assert that four factors affect the spatial patterns of unemployment: migration decisions and labour force participation decisions, both of which are largely governed by households; location of businesses; and wage flexibility. These forces constantly mould the unemployment landscape.

2.2 Determinants of job searching

All search models are designed to recreate the real-world process of an individual's attempt to find employment in a market with varying characteristics, which cannot possibly all be accounted for. However, the basic premise of the models is that all individuals seek work, and this generally places them into one of two states: employed or unemployed. Obviously, this is a simplified list of states and is not exhaustive. Individuals seek work while maximising their utility and the expected present value of their lifetime income over some time horizon. This process is costly and is only pursued if the benefits received from wage offers exceed the costs. In these basic search models, the number of firms contacted, the search duration, and the wage offers presented are all random variables. These variables reflect the random nature of the search process and provide a starting point for understanding the rules or determinants that govern an abrupt exit from unemployment as well as predict and interpret changes in the labour supply. Finding the indicators that best decipher this process is important.

Using the SALDRU93 data, Kingdon and Knight (2006) find various differences between the discouraged and searching unemployed. In general, they find that, compared to searchers, the discouraged exhibited lower household incomes, higher unemployment rates per household member

and poorer sanitation and water facilities. They also lived in areas on the urban fringe with little to no access to amenities and services, such as banks, post offices, schools and shops. Kingdon and Knight (2006) show the broad and narrow unemployment gaps for each income decile for one year, and find that the gaps close as incomes rise. The effect of rising household incomes can have two opposite outcomes: higher household incomes allow members to pass some of these funds to their unemployed recipients, encouraging them to begin searching (search financing effect) or leaving them feeling secure enough that they feel that searching is no longer required (income effect).[1] Given the lower mean incomes and higher deprivation among the discouraged, Kingdon and Knight (2006) suggest that discouragement and lack of search are hindered by factors other than income. In other words, the non-searching unemployed have not chosen to be in this state, but the prevailing economic conditions have disheartened them enough to stop actively searching. Kingdon and Knight (2006) use a binary pooled logit model of job searching for all races, and find a positive relationship with the human capital (education) variables and age. However, the benefits of age are found to be at a decreasing rate, with younger and older searchers experiencing the greatest difficulties finding employment (Kingdon and Knight, 2006).

They also use indicators of remoteness to evaluate the cost of job searches. Three proxies for these costs are used: the community's distances from available facilities; erratic, impassable roads; and whether the household was situated in a homeland. All three negatively affect job searching, with the effect of access to facilities being the most significant (Kingdon and Knight, 2006). Wealth and household non-earned income are also positively associated with search behaviour. Kindgon and Knight (2006) suggest that the search financing effect of increased income is much stronger in the South African context given these outcomes. If the income effect was stronger, a negative effect would have been generated for home ownership and income from social grants (non-earned incomes). This further shows that search costs are a significant deterrent to job searching. Kingdon and Knight also use a poverty dummy variable—reflecting whether an individual lives in a household belonging

1 Kingdon and Knight use the term income effect loosely. The income effect is generally referred to when an individual's income changes, as does its net effect on demand for particular goods. Their interpretation suggests that "additional income reduces the chances of search" (Kingdon and Knight, 2006, p.11).

to the poorest income quartile—and find a negative effect. This further confirms the search financing effect (or discouragement) interpretation. Kingdon and Knight (2006) proxy discouragement with the prevailing cluster unemployment rate and find that is has a negative and significant effect, confirming the assertions of prior international studies, which suggest that increased unemployment rates have a dampening effect on search intensity. Labour market networks for unemployed individuals are proxied using two variables: employed household members and members who are periodically absent from the household (Kingdon and Knight, 2006).

Eriksson et al. (2002) discuss the determinants of search efforts, which they deem to be a function primarily of market demand and individual characteristics. They apply data from three Nordic countries to a simple search framework, and separate the search process into two parts based on different mechanics. The first part deals with the decision to search at all and then, if searching is optimal, the second part focuses on search intensity. They assert that analysing these two mechanics as a single decision could be misleading. In their analysis, the decision to search stems from general attitudes toward working, while search intensity is influenced more by the consequences experienced during unemployment. Their paper analyses the effect of unemployment insurance benefits on search intensity among these subgroups while controlling for various individual characteristics to understand the differences between job search strategies and the effect of unemployment rates. Most of the individuals in Eriksson et al.'s (2002) sample are unemployed. Their sample is separated into two sub-samples for each country: the stock and flow (representing the unemployed and those who have exited unemployment, respectively). The study measures search intensity in two ways. First, an index is constructed from 0 to 9. The index is based on a point system, where a point is awarded for each activity (nine in total) that helped an individual to secure future employment in the previous four weeks. Second, reported hours spent searching are used to indicate search duration. They also assume a fixed time cost per time unit of search, which can be viewed as a simple proxy for transport costs. Ultimately, Eriksson et al. (2002) find that individual characteristics are a weak indicator of the search decision. They also find that in two of the Nordic countries, being involved in a labour market programme reduces the propensity to search, while increasing search behaviour in the future. Job

search increases were found for those who had unemployment insurance benefits, or for individuals who became unemployed because of economic reasons or increased education. Education also positively affected job search intensity. Eriksson et al. (2002) assert that an individual's attitude towards work (those who prioritised job searching) was the most important factor affecting the job search decision.

Hinks (2008) uses local unemployment rates, education level and social networks as determinants of job searching. He includes previous work experience, household job security of employed members, household pensioners, poverty and proximity to public transport, which all affect the job search decision to varying degrees. Hinks (2008) analyses these effects by applying a binary logit model to the search. He focuses on the unemployed (justifying this because previous analyses have found large differences between the types and quality of employment, which would overcomplicate the model) and accounts for urban and rural effects.

Using the 2000 KMP Survey, Schoer and Leibbrandt (2006) find that, of those who were currently employed, only a third achieved this status through formal channels. The remaining respondents found employment through social networks. Their paper shows why search theory should look past the simple dichotomous decision to search or not to search. The relationship between an individual's search effectiveness and the feasibility of the type of search method play an important role in optimal search strategy determination (Schoer and Leibbrandt, 2006). It is important to note that, although Schoer and Leibbrandt (2006) assert that the search method is important to the decision making process, it is most likely that in standard search models, this decision is hidden within the search cost itself. If the searcher is a value maximising agent, one expects that choice maximises the problem. Schoer and Leibbrandt (2006), similar to Eriksson (2002), divide the search process into two distinct and consecutive processes: identification and acquisition of a job. Each process has its own probability distribution. Job identification relies on individual and household characteristics, while job acquisition relies on employer recruitment methods (Schoer and Leibbrandt, 2006). Schoer and Leibbrandt focus on the process of identification and use a multinomial logit model to assess each of the four search methods: exclusive search, passive search, a combination of exclusive search and non-

searchers. They find that, apart from gender, all other individual characteristics (age, race, marital status, and education) are not important to choosing a search method. However, they find varying results for household characteristics that discourage or encourage certain search methods. In particular, they assert that domestic duties limit the use of strategies other than passive search. Schoer and Leibbrandt (2006) conclude that certain household characteristics can limit a searcher's choice of search methods and ultimately their search decision.

Carolina and Pau (2008) discuss two classifications of the job search literature. They discuss the information gathering approach, which is similar to Stigler's view of searching. In this approach, information required from employers and job seekers is incomplete, therefore both must expend time and resources to gather information regarding wages, vacancies or skills (Carolina and Pau, 2008). This process is costly and time consuming to both parties. They also discuss the trade frictions approach, which views unemployment as a consequence of mismatches (or trade frictions) between employer vacancies and job seeker applications. Carolina and Pau's (2008) model is constructed on the information gathering approach, which uses a cost benefit analysis as the individual aims to find better offers while adhering to a minimum wage. They classify determinants of search as economic resources (i.e., income related factors), individual and household characteristics and institutional factors (i.e., unemployment insurance and a range of welfare benefits). Unemployment duration is another potential search factor that Carolina and Pau (2008) include. However, the ambiguous effect that unemployment duration has on the reservation wage leaves it open to interpretation.

Singell and Lillydahl (1985) investigate the effect of residential decisions on household job locations, focusing on the difference that gender makes to these outcomes. Their estimations are modelled on three equations, which centre on commute times and wage and rent gradients and are examined against a host of socioeconomic and role-related variables for two-earner households in the United States. Separate estimations are generated for men and women as well as for those that had a change in residence and those that had not. They propose that job locations are pivotal for men with regard to residential decisions and use work commute differences between men and women to tease out its effect. They find that job location for men significantly affects household migration behaviour.

Gender and race wage differentials from commuting are also significant. For men, commute times were shorter, but only for part-time work. Singell and Lillydahl (1985) also find significant differences in commute times for the types of jobs dominated by men or women. The former have consistently longer commute times, regardless of whether a man or women filled the position. These commute time differences between men and women accounted for as much as a 10% difference in wage rates (Singel and Lillydahll, 1985).

2.3 Transport costs

Transport costs are at the heart of active searching in South Africa. Those who live in peri-urban or settlement communities mostly use either taxis or trains to travel into urban city centres for work or the prospect of work. Isolating this component of search costs is difficult from both a data and a research standpoint. One of the main difficulties of discussing transport costs on their own is that they are often not explicitly marked as discussion topics in papers that analyse job searching. For the most part, they are dealt with in varying ways to accommodate the different approaches advanced by various authors. Church et al. (2000) express similar concerns, arguing that transport is rarely addressed in the unemployment, search and wage literature, and that most studies proxy these costs with other variables. This almost seems counter intuitive given the impact that mobility can have on individual search prospects. Since most of the costs related to travel are calculated as distances or times, a spatial aspect to this analysis would be useful. A brief discussion of these works follows to give insight into the importance of these costs as well as their treatment.

Rouwendal and Rietvield (1994) use commuting distances of Dutch households as well as personal and labour market characteristics, which are modelled on the theoretical economic theory of search. They use a multivariate analysis of commuting distances and augment standard search theory by including a spatial aspect in the model. This is done simply by including the cost of transport in the optimisation problem. A searcher minimises the cost from the wage offer before evaluating it against their reservation wage. Rouwendal and Rietvield (1994) assume that commuting costs are an increasing function of commuting distance, and hence the variable is spatially distributed by employment. This further implies that the probability distribution function of jobs and distances are

the same (Rouwendal and Rietvield, 1994). They also show that by using the theory of search in this way, the distribution of observed commuting distances can be directly related to the geographical distribution of job opportunities around a household. The commute distances used by Rouwendal and Rietvield (1994) reflect how far members of a household live from their place of work in 10-km increments. This means that they do not use the actual commuting distance, the dependent variable, and manage this problem by using a special case regression model for censored data. Therefore, their dependent variable is a continuous scale where actual values are unobserved. Although their analysis involves mostly employed workers, as well as changes to housing positions, it gives insight into how commute distances affect search decisions. The most significant of the variables was age, which had a negative relationship with commute distance. They assert that this could indicate that older people prefer shorter commutes. They also find that education and hours worked had a positive and significant relationship with commute distance, but no location variables had any impact.

Van Ommeren et al. (1998) explore individual reactions to searching from increased costs and travel time caused by congestion or road pricing. They investigate on-the-job searching as a function of commuting time in the Netherlands, and show that a convex relationship exists. This could translate to increased search intensity as individuals devote more commuting time to finding work. Furthermore, they focus on employed individuals who search for better jobs and housing and find that the expected life value of a current job diminishes as commute time increases.

Van Ommeren and Straaten (2005) produce similar research and base their analysis on the wasteful commuting literature. They assert that this literature is important because it is used to evaluate commuting cost minimisation assumptions. Their paper assesses whether excess commutes, defined as the difference between the average commute and the minimum commute in a given area, are much larger for employees than the self-employed. Their premise is that self-employed individuals should have negligible excess commutes because the arrival rate of vacancies is infinite relative to the job arrival rates of employees. Van Ommeren and Straaten (2005) modify the standard search model by substituting the wage distribution that a worker faces with the distribution of commuting distances. Using this model, they show that excess commutes for the employed group depend on the job arrival

rate. For the self-employed group, the excess commutes depend on the workplace arrival rate, which is much higher, hence, as it approaches infinity, the excess approaches zero. They use distance and time to measure commutes, and assume that the commuting costs are proportional to the commuting distance, which they find renders similar results. Van Ommeren and Straaten (2005) test the ideas proposed by their modified model empirically, finding that commuting times and distances are large among the employees. Furthermore, the excess commute equals the average difference of both groups, which is around 40% to 50%. This is even smaller in urban areas (Van Ommeren and Straaten, 2005). Wasteful commuting is premised on measuring market imperfections that lead individuals to sub-optimally choose residences or work-place locations (Ommeren and Straaten, 2005). Both affect the distance travelled for work and job searches. These costs are important because, in general, it is expected that the discounted value of a residential move closer to work is greater than the discounted flows of reduced commute distances (Ommeren and Straaten, 2005). This shows that individuals who are unable or unwilling to move residences, typically the rural unemployed, suffer even greater search costs. This is important given the high rural unemployment rate that exists in South Africa.

Van der Berg and Gorter (1997) use a structural approach to empirically analyse job search theory for unemployed individuals. They report that much of the literature surrounding job-search behaviour neglects commuting time as a job search characteristic, even though it could be an important determinant of search behaviour. Focusing on commute times would allow studies to move away from the typical one-dimensional reservation wage model and allow jobs with lower wages and shorter commute times, and vice versa, to be considered (Van der Berg and Gorter, 1997). Therefore, an individual's optimal strategy is no longer based solely on the reservation wage hypothesis, but rather on a combination of wages and commute times. This intuitive approach allows individuals to accept low wage offers with short commutes or a job within walking distance, or reject high wage offers with longer commutes, and ultimately determine the effect of commuting costs on job search behaviour. Van der Berg and Gorter (1997) define their parameter of interest, the utility trade-off between wages and commuting time, by using the reservation wage difference between the jobs accepted by two

groups: those with almost no commuting times and those who commute an hour to work. They claim that doing so determines the variable simply and reduces the impact of nuisance factors. They find that individuals living in urban areas are willing to pay substantially more for commutes than their rural counterparts. This may suggest that urban individuals include personal and other leisure activities in their trips or that shorter commutes to and from leisure or work places are waylaid by traffic congestion (Van der Berg and Gorter, 1997). Both these alternatives are unlikely to occur in a rural setting. Van der Berg and Gorter (1997) also find that personal characteristics have an insignificant effect on the relationship and exclude them from their final estimation.

White (1986) attempts to bridge opposing ideological views held by urban and labour economists regarding the commuting patterns of urban individuals. Urban economists contend that individuals in monocentric cities have fixed job locations and are compensated with declining housing prices as they move further from the city centre. Labour economists view the residential location as fixed and commuting compensation is addressed in an individual's wages. White (1986) develops a model that simultaneously combines and determines the locations and their respective compensations as well as accounts for gender differences. He uses a reduced form model to estimate the impact of taste and demographic factors, which differentiate households or workers from each other, on commuting time. He finds that there is a negative relationship between commuting time and residential tenure, suggesting that these households may prefer to find jobs closer to home and accept reduced wages. In particular, female household heads have increased commuting times if they have young children, which may indicate that working women stretch themselves by finding work that is further from the household (White, 1986). Furthermore, he finds that the most important determinants affecting commuting behaviour are individual tastes and demographic factors.

Van Ommeren (1998) uses a logit model to determine search efforts in two different model specifications: commuting time and commuting distance. Both these independent variables are measured on a continuous scale. Their main finding is that both travel time and distance affect job search effort positively and in a convex manner. However, they find that on-the-job work seekers are more willing to pursue employment that requires long commute times than they are when large

distances are involved. Van Ommeren (1998) asserts that the former is a determinant of commuting costs. In contrast, Small and Song (1992) argue that both time and distance are equivalent when determining commuting costs. Van Ommeren (1998) also finds that commuter socioeconomic characteristics do not have a significant effect on commuting times that are devoted to searching. These findings contradict the standard economic literature, which asserts that commuting time is compensated by reduced housing prices and higher wages for locations far from the city centre (Van Ommeren, 1998; White, 1986).

Van Ommeren and Fosgerau (2009) suggest that many studies that focus on the travel time component in their assessment of commute costs do so using commuter mode and route choices. This generally seems appropriate because mode and route choices ultimately have the largest effect on costs, since this leads to the consequential costs of the decision. However, the difficulty of equating different transport alternatives in terms of travel time and cost and the correlation between travel time and cost remains a flaw in these studies, since revealed preferences become inherent in these models (Van Ommeren and Fosgerau, 2009). Revealed preferences are the result of individual actions that reveal their preferences, and can produce results that seem counter intuitive. For example, an individual may say that they really dislike commuting, but then chooses to live a long way from work and commute a long distance, therefore revealing their preference for longer commutes. Van Ommeren and Fosgerau (2009) suggest that this can be avoided by using data that captures subjective responses on hypothetical trips or modes, but inadvertently introduces biases related to hypothetical data.

The cost per trip, according to Van Ommeren and Fosgerau (2009), is easy to determine for those using public transport, since the ticket price can be used as a good indicator of the marginal monetary cost. This is not necessarily true for commuters who drive cars to work because trip costs include other costs related to car ownership. However, Van Ommeren and Fosgerau (2009) estimate commuting costs by using a dynamic job search approach, which includes on-the-job searching and behaviour. Dynamic models are those that model a process over multiple time instances or a process that changes over time. Using this method and including job moving behaviour helps to overcome

some of the strong assumptions of discrete choice base models, one of which is that costs and commute time are only affected by mode choice. This is not a major concern for the South African data because car availability or ownership is low among poor unemployed individuals.

As with many studies, time and distance are often used as proxies for commuting costs, because the available data does not usually include actual commute costs (Van Ommeren and Fosgerau, 2009). Deciding whether to use time or distance to proxy individual commuting costs is largely affected by whether there are large variations in speed—commuting time and distance would be used respectively in each case (Van Ommeren and Fosgerau, 2009). Van Ommeren and Fosgerau argue that time and distance can only be used interchangeably when commuting speed is fixed and constant across the population. Small and Song (1992) argue that there is no difference between either measure. Furthermore, using the commute time is only theoretically justified when commuting speed is endogenously chosen (Van Ommeren and Fosgerau, 2009). This makes sense for individuals who own cars, since their commuting speeds can vary greatly, given the large array of factors that could affect speed, and is not necessarily a good measure for those without cars.

The three most important factors affecting whether an individual chooses to search or not are: the prospects of attaining work with a high unemployment rates, recruitment strategies used by potential employers and job search costs (Kingdon and Knight, 2006). Of these, the last is the most important to this paper because it includes transport costs in its scope. Kingdon and Knight (2006) find that 18% of the discouraged cited commute costs as one of the main reasons for their disinterest in searching for a job. This cost can be a serious impediment to poor people living in remote areas (Kingdon and Knight, 2006). Wilson and Ramphele (1989) demonstrate how large these costs are to rural inhabitants by using KwaNdebele, a former homeland located north east of Pretoria, as an example. Depending on the type of transport used, commutes into Pretoria required two to three hours for a single trip. This means that a daily commute, for searchers or workers, took about six hours and incurred transport costs. Although KwaNdebele has seen some development since then, many KwaNdedele residents still face these enormous obstacles. Given low employment prospects, individuals may find that the burden of their commuting search behaviour is too financially and

emotionally taxing. Kingdon and Knight (2006) use the homeland areas as an example where the benefit to cost ratio is too low, given the poor labour absorption rate and long travelling times. This is probably why the unemployed living in homeland areas commonly rely on word of mouth communication between friends and family who live closer to employment zones to relay potential employment opportunities (Kingdon and Knight, 2006).

Using a theoretical model to disseminate the different commute and search behaviour of whites and non-whites in England, Patacchini and Zenou (2005) find that both groups decreased their searching as the distance to job opportunities increased. Furthermore, the white group searched more intensely for any given commute distance to jobs. This, they assert, is because the white group had easier and cheaper access to employment availability. The overarching conclusion is that for each race there is a negative relationship between distance to jobs and search efforts. Access to a vehicle increases search efforts for all race groups. For this reason, they propose that giving non-whites easier access to cars would go a long way in bridging the search intensity gap between the groups. Patacchini and Zenou (2005) also use the SMH to indicate why ethnic minorities in the United States and Europe suffer high unemployment rates. As the theory states, distance to jobs is the main barrier to employment opportunities for members of this group, who are often positioned in areas that are remote and segregated from urban society. They also investigate the transport mode as a determinant of the search intensity gap that exists between white and non-white groups. In the United States, where these differences are well documented (Patacchini and Zenou, 2005), it is generally accepted that minorities (black workers for instance) mainly use public transport while whites have access to cars. Therefore, blacks are more likely to refuse, or search less, for jobs that require long commutes and would rather focus their search in areas located closer to their neighbourhoods. This is akin to Kingdon and Knight's (2006) benefit to cost ratio: increased commutes decrease the ratio and increase reservation wages. Patacchini and Zenou (2005) also motivate that the use of private transportation by whites gives them advantages in terms of time, cost and information gathering when searching for jobs. Their analysis ties the ideas presented in the SMH and transport mode together by using the Nomenclature of Territorial Units for Statistics (NUTS3) and the English LFS data. They do this by creating two

models: one where whites use cars and the other where non-whites use public transportation. To accommodate the lack of data related to commute times for the unemployed, a proxy is created using the average commute times for employed people living in similar areas (Patacchini and Zenou, 2005). Their results show that, if unemployed workers have poor access to jobs (they are far away in terms of time distance), then they search less than employed people, who have better access, because it takes more time (and is thus more costly) to gather information about jobs. Furthermore, they assert that low cost housing in remote areas decreases the incentive to search because there is a less need for money for housing, which results in unemployed individuals searching less (Patacchini and Zenou, 2005).

Smith and Zenou (2003) create a model to demonstrate an alternative theory to the SMH. By using a search-matching model, they show that the unemployed optimise their long-term unemployment choice and search less. This is supported by showing how increased distance to job centres results in lower search intensities for all groups, thus creating a link between unemployed search intensities and locations. Living close to job opportunities, mainly located in CBDs, provides short-term disincentives in the form of increased rent and lower housing consumption, but long-term benefits through increased search intensity, resulting in an increased probability of reemployment (Smith and Zenou, 2003). Unemployed labour market participants in remote regions benefit from lower rent and higher housing consumption in the short run. However, they suffer in the long run from decreased searching, because of the cost, which in turn diminishes the chances of re-employment (Smith and Zenou, 2003).

Wasmer and Zenou (2002) develop a model centred on search efficiency and how it is negatively affected by distant job access. The model builds on the SMH developed by Kain. Job seekers from remote areas receive less information regarding distant jobs because firms opt for local recruiting methods, making this type of searching very inefficient (Wasmer and Zenou, 2002). They assert that labour force participants residing close to job locations have better access to information pertaining to jobs, and are therefore generally more successful in their search activities. Their framework involves constructing two scenarios: in the first, a segregated city, the unemployed reside far from jobs; and in

the second, an integrated city, unemployed individuals are located close to jobs. They use this to determine that there are small welfare and cost decreases for longer commutes, despite the search inefficiencies in the segregated city. Their theoretical intra-urban search model combines both search and urban models. By including distance in their model, Wasmer and Zenou (2002) deviate from the standard job matching models and strongly link unemployment with location. This negative relationship between distance and employment exists because increased distance has a large, negative impact on search behaviour (Wasmer and Zenou, 2002). Linked to this, is the information loss associated with increased distance to employment centres since these unemployed searchers are less likely to find employment adverts. Therefore, it is assumed that both the unemployed and employed prefer to be closer to CBDs because the employed want to reduce commute costs, while the unemployed want to increase their probability of getting a job (Wasmer and Zenou, 2002).

Patacchini and Zenou (2005) investigate the differences between white and non-white groups in England based on their search intensity. They do this by comparing the two groups in terms of transportation modes (private and public transport, respectively), while keeping the location from job centres (or the distance) the same for the two groups. Assuming that private transport is quicker, Patacchini and Zenou (2005) assert that commuting time (distance) would be lower, resulting in lower information-gathering costs for the white group. Their model is justified with empirical evidence based on a pooled fixed effects model using the NUTS3 English panel data. Inter-race differences are estimated using an Oaxaca decomposition. Patacchini and Zenou (2005) also find that for the low skilled unemployed, both access to jobs and cars have a larger negative impact on their search intensity. This can be attributed to their distant residential locations and the inability of the wages they attain through possible employment to compensate them for long commutes. This is particularly true for the non-white group, who generally live further away from business centres. They conclude that there is evidence to show that the differences in search intensities between whites and non-whites can be attributed to job access (time distance to jobs) and car ownership differences.

2.4 Data

Although there are a number of labour market surveys in South Africa, very few explicitly collect data on transport costs. This research uses two recent data sets which do. These datasets are the LMES and the NIDS.

Measuring transport costs seems easy—one simply has to ask "How much does it cost to travel to and from places you look for work or to and from work?" This would be the ideal question for this paper. If this question was asked, it would be a rather straightforward method of capturing those individuals who are searching and unemployed and finding out whether a negative relationship exists between the two variables. The NIDS data provides for this from a question posed to the unemployed regarding the search costs. The LMES data captured transport costs only for employed individuals which meant constructing a transport cost for the unemployed. This paper uses other control variables, such as those utilised by Hinks (2008) and Kingdon and Knight (2006), and controls for the endogeneity created by the individual choice of transport mode, transport cost and residential location. The analysis is conducted at a cluster level, because doing so allows the probability of search across all clusters to be more generalized. Transport data relevant to this paper comes from the NIDS and the LMES data sets. Both data sets are similar in that many of the questions in each survey are almost identical. This enriches the analysis by adding a degree of comparability across surveys.

Table 2.1 compares the LMES and NIDS data sets and classifies respondents by their labour market status. The four statuses are constructed based on standard classifications. The employed are those people who stated that they were currently employed, had a secondary occupation, were self-employed, were employed casually, were involved in personal agriculture or assisted others in their businesses. People who are not economically active (NEA) were not working and had not desired work in the last four weeks. Two unemployment categories are included: one that includes those who are discouraged (people who desire work, but have done nothing to look for work) and a strict definition which simply states that individuals are unemployed if they are actively seeking employment. The LMES has various questions that allow for two separate definitions of these states. One question asked *"What activity currently takes up most of your time?"* Respondents answered this

question by selecting one of six responses. This question was used to identify whether respondents were employed , NEA, discouraged or unemployed. However, other questions were also included in the interview to allow for a more precise definition of employment states. These questions were used to construct the last column of results in Table 2.1 (referred to as LMES controlled). The two data sets vary greatly both in size and in their relative composition. The NIDS includes a high proportion of NEA individuals (43%), while the LMES and LMES controlled are only composed of approximately 12% and 24% NEA individuals, respectively. The narrow definitions, which include people who are unemployed and discouraged or searching, show that the NIDS, LMES and LMES controlled include 18%, 60% and 34% unemployed respondents. This suggests that the LMES respondents experience almost two and half times more unemployment than the national average, which was around 34.9% in 2011 (Development Policy Research Unit (DPRU), 2012). It should be noted that, although the surveys use the same classification definitions, it was difficult to perfectly apply the definitions across the data sets, even though the LMES and NIDS asked questions which were used to classify respondent employment states. These are standard classifications and future surveys should define slightly stricter versions of these variables, particularly for those who are discouraged and strictly unemployed. The LMES data only includes individuals from 20 to 24 years old. This is probably why the LMES shows larger unemployment rates because of the prevalence of youth unemployment in South Africa.

Table 2-1: Employment status according to the NIDS and LMES

Status	NIDS		LMES 2011		LMES (controlled)	
	No.	Col %	No.	Col %	No.	Col %
NEA	6663	43	165	12	333	24
Discouraged	976	6	216	16	96	7
Unemployed strict	1838	12	527	38	369	27
Employed	6012	39	459	34	596	42
Total	15489	100	1367	100	1367	100

Several questions in both surveys pertain to transport as a whole. These include questions concerning both transport modes and costs. This paper uses the NIDS measure of household expenditure on transport in the last 30 days, as stated by the respondent at the time of the interview. It also uses a

measure of the amount spent on transport in the last 30 days by individuals. These variables are measured in Rands. The LMES only provides transport costs for employed respondents, but measures were taken to apply these costs to the entire sample. Data from Google Maps is also included to augment the LMES data and provide another proxy for transport costs. The specifics of the transport questions and the related data are discussed in the following chapters.

3 Chapter 3

3.1 Introduction

This chapter examines the relationship between transport costs and searching. As the theoretical models reviewed in Chapter 2 show, search costs are an important determinant of equilibrium labour market outcomes. Individuals need money to travel and if the expected benefits of not searching outweigh the costs of searching, then it makes no sense to continue searching for a job. To investigate this in more detail, this chapter focuses on replicating the work of Kingdon and Knight (2006), who apply a pooled binary logit model to the decision to search whilst unemployed. Unlike Kindon and Knight, this study uses an actual measure of transport costs. This paper uses many of the variables used in Kingdon and Knight's paper, allowing the results to be compared. Other transport and related variables are also included in the regression.

Transport costs are an integral part of searching for a job and commuting to work. They are even more crucial for unskilled job seekers who live on the edge of urban society, where access to facilities, which aid job searches, is limited or non-existent. This includes (but is not limited to) the Internet, labour market networks and information linked to potential jobs. As job seekers try to eliminate some of the costs of searching and travelling, they face even more problems. Transport data used in this chapter is from the NIDS wave 1 data, which was completed in 2009.

3.2 Transport costs and job searching

Much of the literature on the relationship between search and commute costs has been discussed in the previous chapter. However, several other studies deal with the general topic of spatial mismatch and employment, which examine the spatial problems associated with transport costs and searching. Miller and Ong (2003) suggest that, to overcome the spatial problems associated with distances to and from work, job seekers need greater access to vehicle purchasing, since personal vehicle use has many benefits and improves labour market outcomes. Their paper employs a weighted least square, weighted two-stage least square and ordinary least square (OLS) regression to estimate the impact of households without cars on both employment and unemployment. A weighted two-stage regression is used to eliminate some of the endogeneity problems faced in their study. Miller and Ong (2003) use

their findings to justify the ineffectiveness of transit routes, or other similar transport policy measures, to overcome transport barriers to employment, since they find that access to a vehicle independently improves employment outcomes. Robins (1981) asserts that, in small cities, jobs are easily accessible by all residents who have access to cars. In the South African context, this does not suggest that the best way to improve the mobility of poor unemployed individuals is to give them better access to cars. However, it does serve to show that increased mobility might narrow the unemployment gap experienced by these individuals.

The SMH, which was first described by Kain in 1968, proposes that black American workers live in segregated zones (city centres), which are separated from major employment areas (suburbs). This meant that they are poorly connected to major centres of growth (Kain, 1968). Gibillon et al. (2007) examine the SMH in detail and present empirical evidence supporting the theory. Although the SMH generally refers to black people who are minorities (which is not the case in South Africa), it indicates the disparities that people on the urban fringe suffer when searching for jobs in urban hubs. Job opportunities available are primarily city-centre based, with many poor black people living well outside of this growth centre. Gibillon et al. (2007) suggest at least seven underlying factors that harm job opportunities because of distance. Three are relevant to this paper: refusal of jobs by workers due to long commutes; the greater the distance to jobs, the less efficient the job seeker becomes in their job search; and worker location and distance from jobs may increase search costs

Wittenberg (2001) states that search costs are particularly important when they can be reduced through social network signalling effects—individuals who are well positioned to access employer and job suitability information can reduce their search costs per job match. Although Wittenberg focuses on access to employment through household and neighbourhood networks, he gives insight into the assessment of the potential effect of the search costs and reservation wages of these individuals.

Holzer and Ihlanfeldt (1996) use a linear probability model (LPM) to assess whether the proximity of firms to both residential locations and public transit points affect black employment and wages in

Atlanta, Boston, Detroit and Los Angeles. They use firm level data and find that the probability of employment increases the closer the employer is to the black employee's public transport point and residence. Their findings are consistent with the SMH and also point out that, although black employees live closer to employers, on average, their cost per mile to access those jobs is higher than whites who live in outlying suburban areas. Holzer and Ihlanfeldt (1996) further assert that policies focused on transportation improvement must be designed to improve access to higher wage jobs. This is somewhat in contrast to Miller and Ong (2003) who suggest that vehicle ownership independently improves employment outcomes.

3.3 Problems faced and resolved

One of the main improvements of this research over existing research is that it incorporates commute costs into the analysis of search decisions of people who are unemployed. The NIDS dataset asked people who were unemployed and searching, but not discouraged, what their transport costs were. This means that including the variable in a logit regression results in zero observations because people who are not searching have no search costs. Therefore, the logit estimation has no values to base a comparison on. One way to solve this problem is to create a "no search transport cost" variable for individuals who are not searching, and assign it a value of zero. Although this circumvents the problem, it does not eliminate the issue of almost perfect correlation between the dependent variable (whether the unemployed search or not) and the transport costs of searching. Including these measures in the regression also creates endogeneity issues. This is because rising transport costs may be positively correlated with increased searching, while increased searching creates higher transport costs. Various ways of controlling for this endogeneity are listed below.

- Creating a cluster wide average of the transport costs and including this as an independent variable. The NIDS respondents were sampled from 400 primary selection units (PSUs) or clusters. By creating an average transport search cost for each individual that is based on only the individuals in that cluster, it is possible to create a transport search proxy for that cluster. This proxy can be used instead of the reported transport cost in the regression.

- The NIDS asks respondents how much they spend (in rands) on transport to work. This can be used as a proxy for transport costs to and from places of work. This method assumes that individuals are searching in areas where people are working. The individual search costs can then be compared, by cluster, against the commute costs.

Creating a cluster wide average transport cost was chosen instead of other methods because it is easy to construct and it also removes some endogeneity associated with commute costs. However, some other issues still remain. Although a negative relationship usually exists between transport costs and search behaviour, the opposite can also be true. Miller and Ong (2003) describe this as a simultaneity issue, because the hypothesised relationship can exist in the opposite direction. For instance, transport costs may negatively affect labour market outcomes or unemployment could affect job seekers' abilities to access transport (Miller and Ong, 2003). Intuitively, it is expected that commute costs rise as people begin searching, as this forms part of their search costs. This poses a problem when trying to apply a regression analysis to the impact of rising transport costs on searching, because this paper argues that the opposite is true. This is because, with all else equal, increased searching increases expenditure. Generally we expect to find that individuals with lower transport costs search more or that increased commute costs reduce searching (the negative relationship). However, it is also expected that individual search behaviour is governed by funds available for transport and proximity to potential jobs. Given two people with similar characteristics who are unemployed and searching and have a limited budget, it is reasonable to expect that the person who lives closer to job hotspots would search more. In this scenario, we expect to find a negative relationship because the same amount of funds result in more searching. This assumes that proximity to search destination is measureable and that individuals in different clusters facing similar transport costs adjust their search frequency based on their proximity to potential employment.

Ideally, panel data is more conducive to this type of analysis because it allows change to be measured with each wave. However, given the unpredictable nature of employment in South Africa, it is unlikely that much benefit would come from discounting the value associated with job searches or employment. For instance, migration to a new location could lower respondent commute costs, which

would affect their decision to search for work. Comparisons with respondents who did not migrate and had high transport costs with discontinued search behaviour would reject the null hypothesis presented here. Given these issues, the results in this paper are intended as descriptive.

3.4 Data—The National Income Dynamics Study

The first wave of the NIDS was conducted in 2008. The survey includes four separate questionnaires designed to capture household composition, household income and expenditure, employment and schooling in South Africa. This chapter combines the household and adult data sets because these two data sets were the most appropriate for the analysis.[2] The household questionnaire was administered to the oldest woman or another household member with knowledge of the living arrangements of the household, while the adult questionnaire was targeted at individuals in the household who were 15 years or older. The NIDS 2008 data set comprises 7305 unique households and is divided into a child and adult component. There are approximately 16885 adult observations. According to the metadata file provided, 1246 of the respondents refused to participate in a large portion of the questionnaire and were therefore excluded from the sample. The NIDS data set comprises three waves of data, with two subsequent waves being conducted in 2010 and 2012. These datasets were not publically available when this study was conducted.

Table 3.1 provides a broad outline of the employment statuses, separated by gender, for the NIDS dataset. In total there are 15489 observations. All variables were constructed using the NIDS metadata provided. Four employment categories were constructed: employed, NEA, discouraged and unemployed. Individuals were classified as employed if they were currently employed, currently had a second job, were currently self-employed, had been paid for casual employment in the last 30 days, conducted personal agriculture activities in the last 30 days or assisted others with business activities. NEA individuals were those who were not employed and had no desire to work in the last four days. The discouraged were unemployed and had a desire to work in the last four days, but had done nothing to search for work or assist others with business activities. Those who were not employed, but had attempted to search for work or assist others with business activities were considered

[2] The child questionnaire is not relevant to this analysis and the proxy questionnaire was only administered when it was not possible to interview the relevant knowledgeable household member (an adult) in person.

unemployed. Women account for 60% of the total sample and they also make up a larger percentage of the various unemployed states. These states are: NEA, broad unemployment and strict unemployment. Surprisingly, almost the same number of men and women are employed. Since this chapter discusses with how transport costs affect job searching, it makes sense to only include unemployed individuals who did something to search for work. The dataset includes 1838 of these individuals, with a skewed distribution of women, who account for 64% of these respondents.

Table 3-1: NIDS—Employment status by gender

Status	Male		Female		Total
	No.	Row %	No.	Row %	No.
NEA	2315	35	4348	65	6663
Discouraged	240	25	736	75	976
Unemployed strict	669	36	1169	64	1838
Employed	2993	50	3019	50	6012
Total	6217	40	9272	60	15489

The accompanying density graph (Fig. 3.1) shows the distribution of household income by gender. For the most part, men and women seem to share an even distribution of incomes, with most individuals earning between R1500 and R2000. The density graph in Figure 3.2 shows the proportion of the mean monthly household income across the four different employment states. Of particular interest are people who are strictly unemployed who share a very small household income. Several of these households appear to have enough income to sustain an unemployed resident's search costs. The graph also shows that most households in South Africa have an income of less than R5000 a month, regardless of their employment status. This might suggest that government transfers make up a considerable portion of the income of many South African households. Figure 3.3 illustrates this point, reflected by the large proportion of low income earners receiving more income from government transfers (blue line) than from other sources.

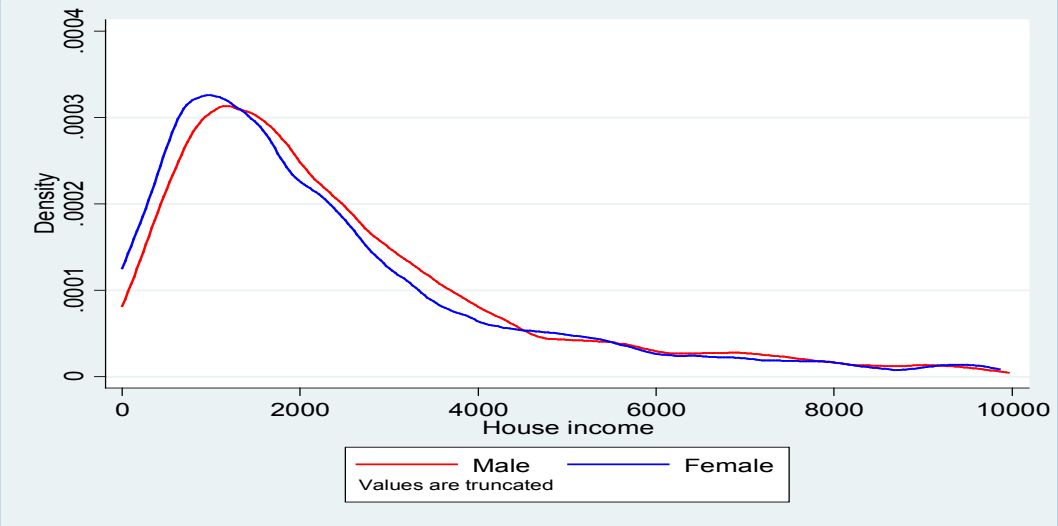


Figure 3-1: Household income distribution by gender

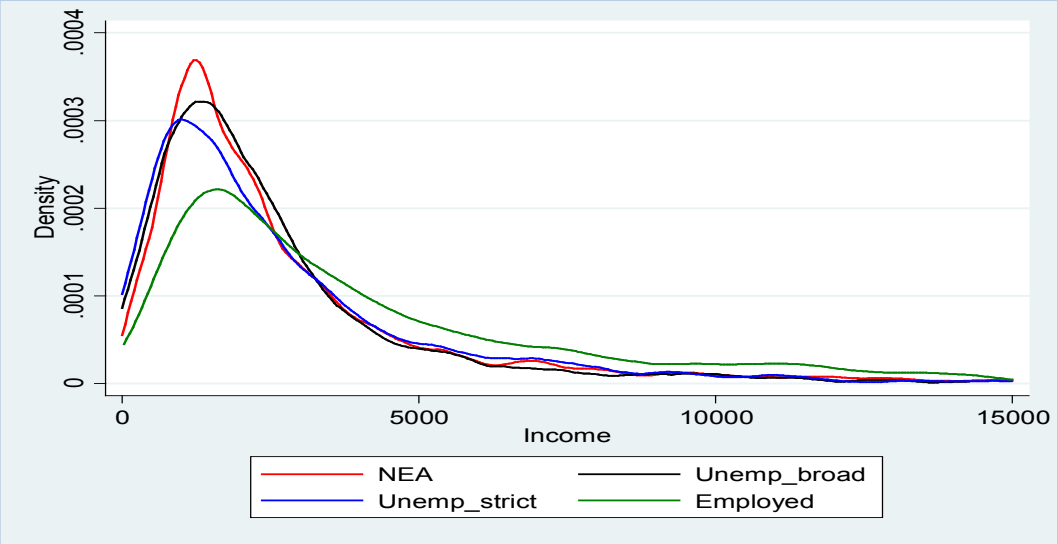


Figure 3-2: Household income distribution by employment status

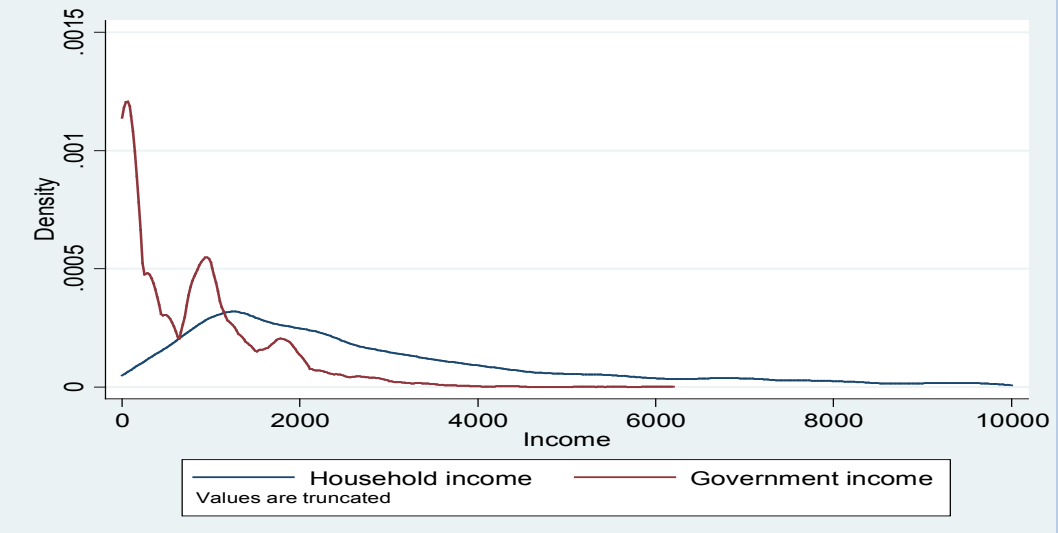


Figure 3-3: Household income distribution (including government income) by gender

3.5 Methodology

A binary logit model of searching among the unemployed is applied to the NIDS 2008 wave 1 dataset. This is done using unemployed searchers and discouraged as the dependent variable together with independent variables that are considered good criteria for the decision to actively or passively search. Many of these dependent variables are comparable with those used by Kingdon and Knight (2006).

3.5.1 Hypothesis

A microeconomic framework is used to analyse the hypothesis that transport costs do not increase the probability of searching. This chapter aims to identify whether people with lower transport costs search more, and uses cross-sectional data to do so. Although these costs can affect both the unemployed and employed, this chapter focuses on the unemployed, since it is assumed that transport costs are a bigger barrier to searching for the unemployed.

3.6 Descriptive analysis (econometric specification)

3.6.1 Dependent variable

A total of 2814 observations exist for unemployed searchers and the discouraged. These observations were used to construct the dependent dummy variable. An individual is considered unemployed and searching (strict definition) if they were unemployed and showed a desire to work as well as some concerted effort to find work (1838 observations). The discouraged group consists of people who were unemployed and showed a desire to work, but did nothing to find work (976 observations). Another way to distinguish these two groups is that their members engage in active and passive searching, respectively (Kingdon and Knight, 2006). Discouraged people search passively for a number of reasons, but this usually occurs when the cost and effort required by active search methods produce negligible benefits. The individuals who did search used various search approaches, the most popular of which was enquiring at workplaces and other places of work, as shown in Table 3.2. Almost 20% of respondents sought assistance from relatives or friends, which might suggest that the costs of more aggressive strategies are less beneficial. Table 3.3 shows the distribution of unemployed searchers and the discouraged. It shows that the largest group came from Kwazulu-Natal (26.5%). There are more searchers than non-searchers, and the searchers make up approximately 65% of the sample. The provinces with the most discouraged people are the Western Cape, Northern Cape,

Kwazulu-Natal and North West, with discouragement levels above 40%. It is interesting to note that Kwazulu-Natal and the North West are former Bantustan territories, which might explain the higher proportion of discouragement in these areas.

Table 3-2: NIDS—reported search strategies

Registered at an employment agency	236	12.8
Enquired at workplaces, farms, or factories, or called on other possible employers	683	37
Placed advertisement(s)	133	7.2
Answered advertisement(s)	208	11.3
Searched through job advertisement(s) on the Internet	65	3.5
Sought assistance from relatives or friends	352	19.1
Looked for land, buildings, or equipment or applied for a permit to start a business or farm	33	1.8
Waited on the side of the road	100	5.4
Sought financial assistance to start a business	5	0.3
Other	32	1.7
Total	1847	100

Table 3-3: NIDS—unemployment status by province

Province of Psu	Discouraged			Searching			Total	Total
	No.	Col%	Row%	No.	Col%	Row%	No.	Col%
Western Cape	109	11.2	**42.1**	150	8.2	57.9	259	9.2
Eastern Cape	86	8.8	25.4	253	13.8	74.6	339	12
Northern Cape	84	8.6	**40.4**	124	6.7	59.6	208	7.4
Free State	61	6.3	26.5	169	9.2	73.5	230	8.2
KwaZulu-Natal	354	36.3	**47.4**	393	21.4	52.6	747	26.5
North West	142	14.5	**43.7**	183	10	56.3	325	11.5
Gauteng	72	7.4	24.6	221	12	75.4	293	10.4
Mpumalanga	41	4.2	21.1	153	8.3	78.9	194	6.9
Limpopo	27	2.8	12.3	192	10.4	87.7	219	7.8
Total	976	100	34.7	1838	100	65.3	2814	100

3.6.2 Independent variables

This section discusses the central covariate in this analysis—transport costs. There are a number of questions in the NIDS pertaining to transport as a whole. These include questions concerning both transport modes and costs. This chapter focuses on the cost rather than the mode of transport, since I argue that cost is a more important determinant of outcomes. As mentioned in Chapter 2, studies have often used time and distance as proxies for commute costs. This is shown by Patacchini and Zenou (2005), Smith and Zenou (2003) and Wasmer and Zenou (2002), who find negative relationships between distance and searching. In contrast, Van Ommeren (1998) finds a positive relationship

between time and searching, which is often considered equivalent to distance in terms of transport costs (see White (1998)). However, Van Ommeren finds a convex relationship, meaning that at some turning point, increased commute times result in less searching. The NIDS dataset does not contain time and distance measures for commutes, therefore these approaches could not be investigated.

The survey includes a measure of individual expenditure on transport in the last 30 days, as given by respondents who answered "*How much did you spend last month on transport to and from this job?*" at the time of the interview. The amount spent on transport in the last seven days is also captured for individuals who answered *"How much did you spend on travel costs associated with looking for work last week?"* These items are measured in rands. Employed individuals answered the first question and the unemployed answered the second. Table 3.4 shows the differences in transport costs between the two groups for each province. The first values (indicated by the *) in each row for each province are the weekly values for transport costs incurred by the unemployed. These values suggest that commute costs for employed individuals are greater than the search costs incurred by the unemployed in every province. Many factors may create this difference, but the irregularity of searching probably accounts for most of the difference. People probably travel to a place of work more consistently than they travel to search for work. Unemployed individuals might only begin searching when an opportunity presents itself or if they hear of a particular job opening. Table 3.4 shows the zero median values for three provinces (Western Cape, Eastern Cape, and KwaZulu-Natal), indicating that a combined total of 391 observations have zero search commute costs. Of the employed commute costs, 1616 observations have zero search commute costs. This suggests that using the mean values to construct the cluster wide variables is likely to be the best option.

Table 3-4: NIDS—comparison of weekly commute costs between the employed and unemployed by province (unit: rands)

Province	Week/Month	Mean	Median	Std.Dev.	Min	Max	Obs.
Western Cape	Weekly	43.21*	**0**	85.21	0	500	145
	Monthly	166.69	**0**	333.20	0	2500	691
Eastern Cape	Weekly	107.98*	40	434.28	0	6000	246
	Monthly	206.26	150	344.49	0	4000	286
Northern Cape	Weekly	54.376*	**0**	253.28	0	2500	117
	Monthly	157.22	**0**	371.93	0	3000	269
Free State	Weekly	92.33*	30	247.25	0	2800	162
	Monthly	172.85	98	254.55	0	1800	234
KwaZulu-Natal	Weekly	96.91*	45	225.58	0	3000	376
	Monthly	183.00	**0**	481.06	0	8000	639
North West	Weekly	137.23*	10	654.54	0	7000	174
	Monthly	271.36	200	411.47	0	4000	289
Gauteng	Weekly	89.37*	31	149.01	0	1000	214
	Monthly	432.35	250	838.21	0	10000	582
Mpumalanga	Weekly	81.363*	5	200.72	0	2000	146
	Monthly	264.97	99.5	500.22	0	4000	272
Limpopo	Weekly	88.34*	30	290.27	0	3333	188
	Monthly	277.92	135	492.37	0	3500	192
Totals	Weekly	91.68*	20	323.59	0	7000	1768
	Monthly	240.11	40	514.09	0	10000	3454

The final transport cost was constructed by using the weekly search costs reported by unemployed respondents. Missing and negative values were removed. This transport cost was then averaged across a cluster so that individuals in each cluster had the same cluster wide transport cost. This was done at the mean and median level. These values were then divided by five to obtain an estimate of the daily commute costs and make the values easier to analyse. Table 3.5 summarises the 25th, 50th and 75th percentiles for transport costs for unemployed and discouraged people. These values reflect what moving from the bottom quarter to the top quarter of the distribution is like. The table shows a steady increase in transport costs from the bottom to the top of the distribution.

Table 3-5: NIDS—Difference between the bottom and top of the transport cost distribution

Variable	Mean	Median	Stv.Dev	Min	Max	Obs.
25th percentile	4.34	0	10.88	0	200	2772
50th percentile	10.79	5.2	19.79	0	250	2772
75th percentile	25.44	16	42.16	0	500	2772

Evaluating the variation between clusters would be useful. However, the NIDS data contains 400 clusters, making it difficult to report values at the cluster level. Collapsing the transport costs by cluster made it possible to find the number of clusters with no transport costs or missing values, while also making it possible to discover if cluster wide transport costs should be constructed using the mean or median values. When using mean values, 48 and 68 observations (clusters) were missing or had zero values for transport costs, respectively. When using median values, 48 and 139 observations were missing or had zero values for transport costs, respectively. This suggests that using the mean values are a better indicator of transport costs because they include fewer zero values. Table 3.6 summarises the number of individuals contributing to each cluster at the provincial level—constructing a table with all 400 clusters would be unmanageable. The last column lists the total number of clusters in which individuals incurred search costs.

Table 3-6: NIDS— summary of the number individuals contributing to cluster wide transport costs

Province	Mean	Median	Std.Dev	Min	Max	Obs.	Clusters
Western Cape	5.25	5	2.79	1	10	145	42
Eastern Cape	8.54	7	5.49	1	21	246	48
Northern Cape	6.61	6	3.22	1	14	117	23
Free State	9.78	9.5	4.73	1	17	162	24
KwaZulu-Natal	7.75	7	4.74	1	18	376	81
North West	9.52	7	6.69	1	24	174	31
Gauteng	6.63	6	2.66	1	13	214	42
Mpumalanga	8.29	7	4.73	1	18	146	26
Limpopo	8.29	8	4.65	1	19	188	35
Total	7.91	7	4.82	1	24	1768	352*

**Excludes the 48 missing clusters*

It is also possible to compare the variation within and between clusters to see if there are evident differences between the clusters. Table 3.7 summaries the differences across clusters for both mean and median cluster wide transport costs. Since both variables were created at the cluster level, there is no variation within them. However, there is a lot of variation between clusters.

Table 3-7: Summary statistics for median daily transport costs across treatment groups

Variable		Mean	Std. Dev.	Min	Max	Observations
m_dcc	overall	20.18	32.81	0.00	250.00	N = 2772
md_dcc		10.79	19.79	0.00	250.00	N = 2772
	between		32.10	0.00	250.00	n = 352
			24.40	0.00	250.00	n = 352
	within		0.00	20.18	20.18	T-bar = 7.875
			0.00	10.79	10.79	T-bar = 7.875

Outliers in the data set suggest that using the median values as a comparative measure would be more appropriate. However, because many median values per cluster were zero it is probably better to use the mean. These values were converted to a daily commute cost, which is easier to understand. The mean daily commute cost for searching varied greatly by province and ranged between R0 and R160, as shown in Table 3.8. This suggests that for most people, searching involved a combination of walking and vehicle transport. Figure 3.4 shows the density of these costs. It indicates that most individuals spend between R0 and R60 per day. Figure 3.5 shows the density distribution of median daily commute costs for each province. All provinces follow the same pattern, with between R0 and R10 spent on commute costs, except for Kwazulu-Natal, which contributes to the large variations in these values, as shown in Table 3.8 and Figure 3.5. This is not surprising considering the dispersed nature of townships in Kwazulu-Natal. Figures 3.6a and 3.6b show the expected negative relationship between mean and median commute costs for searchers. The two figures are very similar, with Figure 3.6a (mean values) showing a tighter grouping along the estimation relationship. This suggests that it would be slightly better to use mean values. However, both measures are used in the final estimations to compare any differences that may exist between the measures.

Table 3-8: NIDS—mean daily commute costs by province

Province	Mean	Median	Stv.Dev	Min	Max	Obs.
Western Cape	4.028395	0	8.411782	0	100	243
Eastern Cape	9.187915	6	14.51888	0	200	331
Northern Cape	10.08019	0	45.11999	0	250	207
Free State	9.134222	8	10.69861	0	60	225
KwaZulu-Natal	13.56129	10	17.29569	0	160	744
North West	16.18395	4	24.92595	0	100	324
Gauteng	11.31512	7.2	14.32193	0	80	291
Mpumalanga	5.430526	3	6.520989	0	20	190
Limpopo	9.667281	6	11.67178	0	45	217
Total	10.79271	5.2	19.78981	0	250	2772

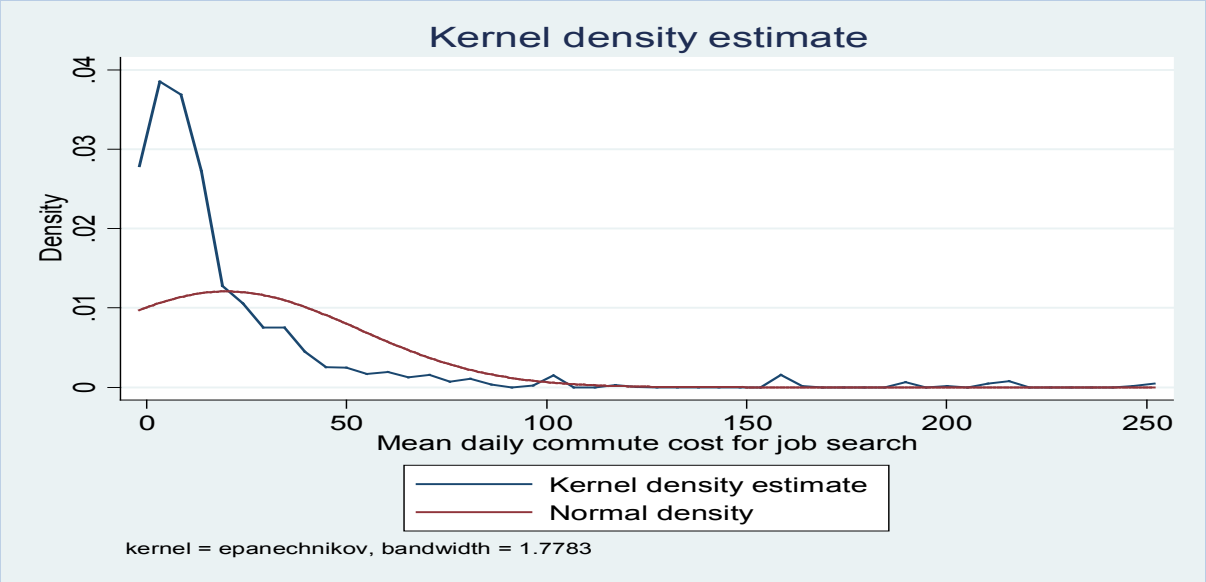


Figure 3-4: Distribution of mean daily commute costs

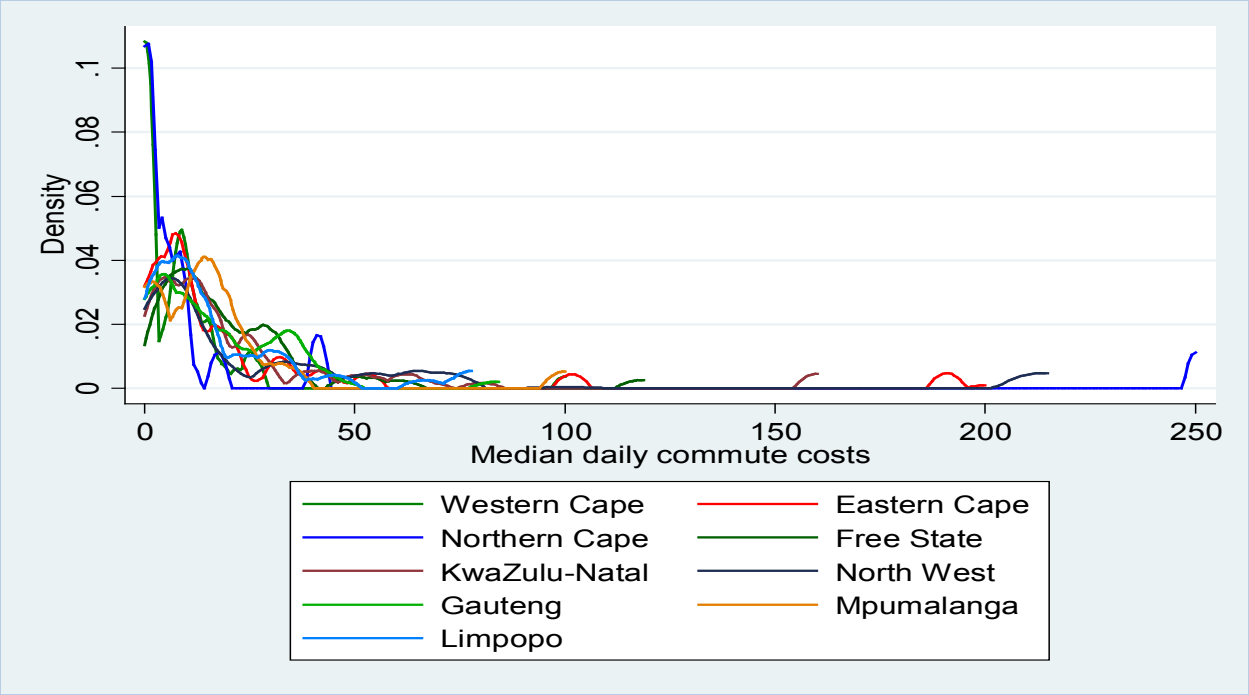


Figure 3-5: Mean daily commute costs by province

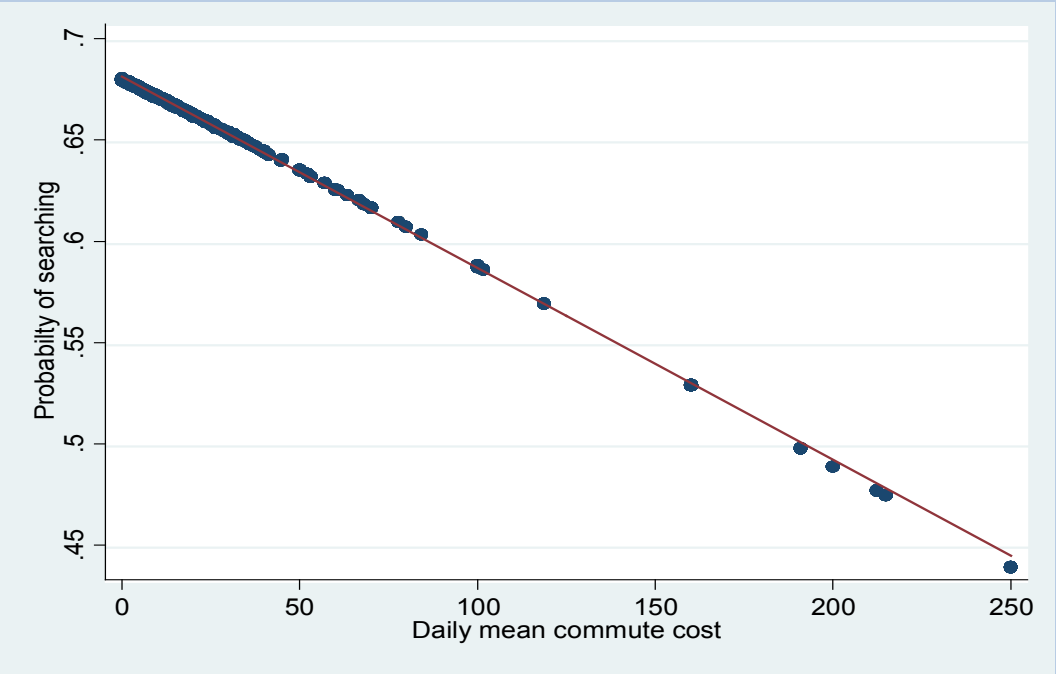


Figure 3-6(a): Observed and fitted logits by mean commute costs

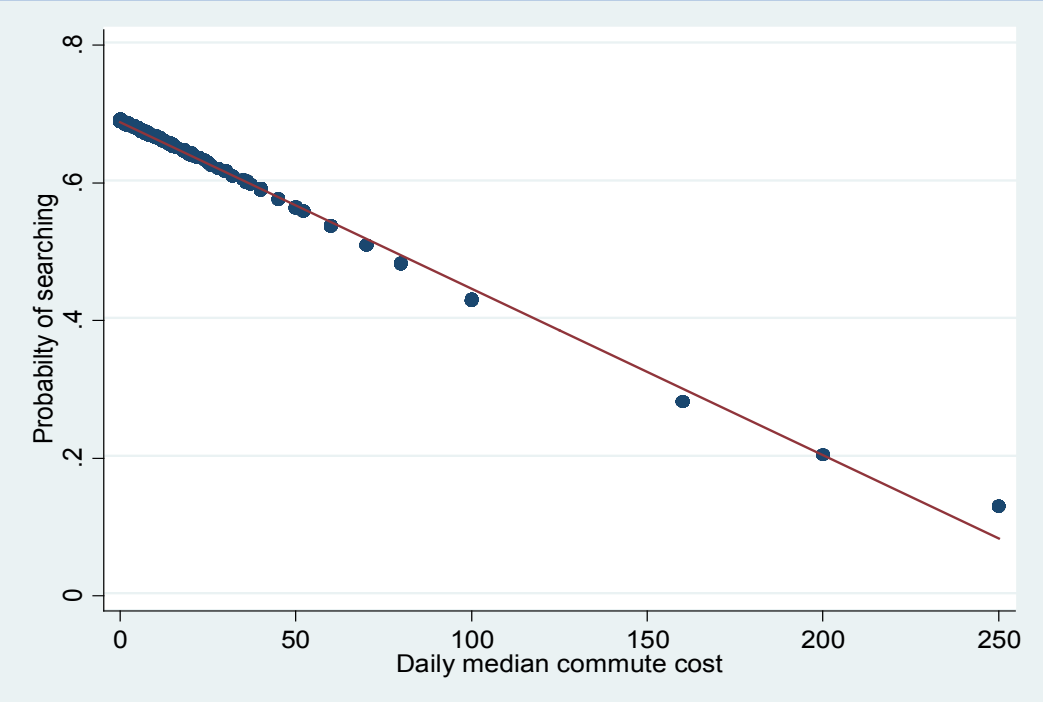


Figure 3-6(b): Observed and fitted logits by median commute costs

This section discusses the other covariates used in the analysis. In order to increase the stability and significance of estimates, it was ensured that most of the covariates are dummy variables. This also serves to make interpreting the coefficients easier. Personal traits and individual characteristics, similar to those found by Kingdon and Knight (2006), are included in this analysis as dummy

variables. These include age, age squared, gender, marital status and whether the individual is the head of the household. Table 3.9 summarises all the variables that are used in the analysis. These figures only reflect the unemployed and discouraged portion of the dataset.

Table 3-9: Summary statistics for all variables

Variable	Mean	Median	Stv.Dev	Min	Max	Obs.
Mean daily commute costs for job search	20.18	10.80	32.81	0	250	2,772
Median daily commute costs for job search	10.79	5.2	19.79	0	250	2,772
Age	31.74	29	11.46	15	90	2,811
Age squared	1,139	841	876.7	225	8,100	2,811
Gender	0.677	1	0.468	0	1	2,814
Married	0.250	0	0.433	0	1	2,347
Head of Household	0.236	0	0.424	0	1	2,814
Primary school	0.852	1	0.356	0	1	2,803
High school	0.334	0	0.472	0	1	2,803
Tertiary education	0.0832	0	0.276	0	1	2,609
Household government income	701.0	420	878.4	0	22,100	2,814
Lives in an owned house	0.830	1	0.376	0	1	2,813
Number of employed members in household	0.613	0	0.822	0	8	2,814
Urban formal	0.421	0	0.494	0	1	2,814
Urban informal	0.0928	0	0.290	0	1	2,814
Rural formal	0.0675	0	0.251	0	1	2,814
Toilet not shared by other households	0.165	0	0.371	0	1	2,560
Cluster unemployment rate	0.216	0.214	0.0855	0.0192	1	2,814
Constant						
Observations	1,996					

Age is a continuous variable, calculated from the NIDS variables provided. It is expected to be positively correlated with searching, because as individuals mature, the responsibility shifts to them to provide for the household. Kingdon and Knight (2006) and Hinks (2008) find that the relationship between these variables is positive. Age squared, as seen in Kingdon and Knight (2006) and Hinks (2008), is included because it is expected that it may have a non-linear relationship with searching. The sample is skewed towards young people with a mean age of approximately 37 years. A large portion of the population is clustered around 20 to 25 years of age with women making up a larger portion of the youth (see the appendix).

Gender differences may affect searches because women are likely to take jobs that are less physically intensive and closer to home, since they may be responsible for child care. Kingdon and Knight

(2006), Hinks (2008) and Schoer and Leibbrandt (2004) find that women are less likely to search for work. Table A3.7 suggests that, on average, women spend more on transport. However, it seems that men commute more, given that there are almost twice as many observations for men. Because the figures for searching are reported on a weekly basis, this implies that each month women and men spend approximately R492 and R292 on transport, respectively (see the appendix).

Marriage is expected to have a positive relationship with searching, as it can provide emotional security to both parties. However, Kingdon and Knight (2006) do not find this variable to be significant. Because of societal norms, marriage may also increase pressure on men to look for work, while decreasing pressure on women, who are expected to fall pregnant. In contrast, Hinks (2008) suggests a negative relationship between marriage and searching. The same applies to individuals who are household heads—the increased responsibility placed on these individuals could increase their search behaviour.

It is generally expected that individuals who increase their level of education increase their wage demands and search behaviour. Eriksson et al. (2002), Hinks (2008) and Kindgon and Knight (2006) find a significant positive relationship between education and searching. This study includes three dummy variables for education. The items from the NIDS adult survey that were used to attain these variables are: *highest grade completed* and *respondent has successfully completed tertiary education*. The former was used to create primary and high school variables and the latter was used to show whether individuals had attained tertiary education. Of the entire sample, 61% of respondents have completed primary school 30% have completed high school and 9% have completed some form of tertiary education. Each of these three categories consists of approximately 30% discouraged people. This may indicate that increased education reduces discouragement or, put differently, promotes searching.

In order to search, individuals need resources and unemployed individuals have smaller income and wealth pools to access. With more resources, it is likely that individuals will pursue more or better employment opportunities. This does not mean that the opposite is true. For instance, individuals who

are better off may be less inclined to search if they have access to sources of income which do not require work. Measuring unobserved wages or the unmeasured value of being unemployed is difficult. Unemployed individuals do not have wages to post. Mortensen (1986) points out that household income can be used to solve this problem. An unemployed individual may not have their own income, but that does not necessarily mean that funds are unavailable. Using household income from sources beyond the control of the individual (transfers) avoids some of the endogeneity issues associated with individual-reported wages and utility maximisation (White, 1986). This study tests four different household income variables. The first of these measures government transfers and other types of non-earned income, which make a large contribution to the incomes of these types of individuals and their households. The household non-earned income variable includes the following items from the NIDS: *household monthly income from government grants*; *household monthly income from other government sources*; *household monthly income from investments*; and *household monthly income of a capital nature*. Household income solely from the government and other government sources is included in the final analysis because of its exogenous nature. A measure of total household income is included and was calculated by aggregating all forms of income except those of a capital nature, but this was not found to be significant. Kingdon and Knight (2006) and Hinks (2008) include household income in their analysis. Kingdon and Knight (2006) also include government income sources. The unemployed and discouraged sample had similar household incomes, with government transfers being more prevalent in searching households.

The question, "*Does a household member own this dwelling?"* captures the home ownership variable. Home ownership can be a significant determinant of job searching. A mortgage on the property might decrease searching because of reduced resources allocated to searching. Kingdon and Knight (2006) find a positive relationship between home ownership and searching, while Hinks (2008) finds a negative correlation. However, ownership can also indicate individuals who have elevated themselves into more secure economic positions, facilitating their own and other household members' search behaviour. Home ownership may also lock an individual into only searching in the area close to their home.

A single continuous variable is used to indicate labour market links by capturing the number of employed members in the household (Kingdon and Knight, 2006). This is expected to have a positive influence on search behaviour, because of the income flows these individuals bring to a residence. Hinks (2008) proxies labour market links with a trade union member in the household, and finds a positive and significant correlation with job searching. Searching households were found to have more employed members.

Kingdon and Knight (2006) use various indicators of remoteness as proxies for transport costs, which could not be duplicated here. However, three different indicators were captured, namely, the geographical area of the sample cluster, whether a household shared a toilet and the cluster unemployment rate. The geographical area is divided into four different types: rural formal, tribal authority, urban formal and urban informal (dummy variables). Tribal and rural areas are more likely to have longer distances to basic amenities, poorer transport facilities and roads, and general difficulties, which often affect those who live closer to urban society less. Most observations for unemployed searchers and the discouraged come from tribal authority areas and urban formal areas. This makes the correlation between rural areas and searching misleading, as the relationship could easily be positive or negative. It seems as though most of the sampling took place in these areas and an estimation must be conducted to reveal its significance. Shared amenities were captured using the *toilet facility shared with other households* variable. This study assumes that individuals who have to share basic facilities, like a toilet, with other households are less likely to be in a position to search for work. However, sharing facilities might result in cost reductions that allow resource allocation to increase or improve search behaviour. Lastly, the cluster unemployment rate is used to determine if particular clusters are prone to unemployment and decreased search behaviour as well as to measure whether being surrounded by unemployment has any effect on the search behaviour of people residing in the same cluster. Increased cluster unemployment is expected to decrease the search behaviour of all individuals (Aragon et al., 2003). Aragon et al. (2003) point out that when using spatial data, the errors that are generated could be spatially auto-correlated. This means that the unemployment rates of neighbouring communities affect each other. Therefore, job availability, or unavailability, in one

area could strongly influence a neighbouring community's unemployment rates, particularly when there are shocks to the labour market (Aragon et al., 2003).

3.7 Results and analysis

Table 3.10a shows the results of a logit model of search among the unemployed of commute costs using the NIDS 2008 wave 1 data, which are the results with no other covariates. This has been done for both mean and median transport costs to compare the two different measures. Both regressions show that the fit is more than adequate with a chi-squared of 0. It shows that both the mean and median daily commute costs are negatively associated with the probability of searching, which means that the hypothesis (transport costs do not increase the probability of searching) is rejected. Although, the log coefficients cannot be directly interpreted, they suggest that the magnitude of the median log-odds units would be greater. This is seen in Table 3.10b, where the probabilities of each incremental change in the transport costs were generated starting at their mean values—R20.18 and R10.79, respectively. There is approximately a 66% chance of searching at the mean value for both covariates. To see if the search probabilities decrease as commute costs rise, the predicted changes for each incremental change in the mean values are shown in Table 3.10b. For both `covariates, the search probability decreases as daily commute costs rise, which is reflected in the mean value increase from R30 to R60. It seems that the median daily commute cost has a larger impact on search probabilities, since there is a 53% search probability at the R60 value. The mean daily commute cost at R60 is approximately 62%.

Table 3-10a: Binary logit of search behaviour among the unemployed

	Mean Values		Median values	
VARIABLES	Coefficients unempsearch	se	Coefficients unempsearch	se
Daily commute cost for job search	-0.00399***	(0.00117)	-0.0109***	(0.00223)
Constant	0.756***	(0.0472)	0.794***	(0.0472)
Observations	2,772		2,772	
Chi-squared	0.000729		1.53e-07	

Table 3-10b: Search probabilities at mean values for mean and median daily commute costs

	Mean dcc	R 30	R 40	R60	Median dcc	R 20	R 30	R40	R60
Pr(y=1x):	0.6628	0.654	0.6449	**0.6264**	0.663	0.6403	0.6149	0.5889	**0.5355**

Table 3.11a shows the results when other covariates are introduced into the logit estimation. It shows that the mean and median daily transport costs seem to have a smaller impact on search probabilities. This can be seen in Table 3.11b, which shows that, for both measures of daily commutes costs, the probability of an individual searching, holding all else constant, is around 70% (up from 64%). Furthermore, almost all variables are significant. The results in Table 3.11a also suggests that age is positively related to search, but at a declining rate; that being a man is negatively related to searching; and that having primary and high school education increases search probabilities. These results are similar to Kingdon and Knight's (2006) results, which find that a number of indicators hamper active job searching. Residing in a formal or informal urban area results in increased search probabilities. The same is true for individuals living in an owned house. Mean and median daily commute costs are negatively related to job searching, which suggests that rising transport costs deter job searching. Table 3.11b shows how R10 incremental changes to mean and median daily commute costs (beginning at their mean values of approximately R20 and R12, respectively) affect the probability of searching, holding all else constant. For mean daily commute costs, this shows that an increase of around R20 to R30 decreases the search probability from 70% to 69.39%, which is a small change. The impact is slightly greater when the median daily commute costs are compared. Overall, when commute costs increase to R60, the probability that an individual will continue searching decreases by 10%. These increments show that increases in commute costs decrease the probability of searching. This is true, but is not accurately reflected by a straight linear relationship as the table suggests.

Table 3-11a: Binary logit of search behaviour among the unemployed (with other covariates)

Variables	Coefficients unempsearch	se	Coefficients unempsearch	se
	Mean values		Median values	
daily commute cost for job search	-0.00361**	(0.00150)	-0.00856***	(0.00262)
Age	0.116***	(0.0275)	0.118***	(0.0279)
Age squared	-0.00176***	(0.000375)	-0.00178***	(0.000382)
Gender	-0.696***	(0.115)	-0.695***	(0.115)
Married	0.0714	(0.152)	0.0644	(0.152)
Head of Household	0.193	(0.149)	0.174	(0.150)
Primary school	0.267*	(0.151)	0.273*	(0.152)
High school	0.652***	(0.127)	0.644***	(0.127)
Tertiary education	0.396*	(0.217)	0.421*	(0.218)
Household government income	-0.000136**	(6.44e-05)	-0.000133**	(6.45e-05)
Lives in an owned house	0.302**	(0.149)	0.305**	(0.150)
Number of employed members in household	0.0267	(0.0616)	0.0159	(0.0618)
Urban formal	0.506***	(0.115)	0.511***	(0.115)
Urban informal	0.662***	(0.203)	0.665***	(0.202)
Rural formal	0.0358	(0.230)	-0.00837	(0.230)
Toilet not shared by other households	0.0472	(0.146)	0.0618	(0.146)
Cluster unemployment rate	-0.331	(0.631)	-0.503	(0.632)
Constant	-1.249**	(0.519)	-1.229**	(0.522)
	1,996		1,996	
	0		0	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Base categories are: female (male==0); not married (marstat==0) no household head (hoh==0); no primary school, high school or tertiary education; no home ownership; is a Tribal area; toilet shared with other households. Values are not adjusted by weights

Table 3-11b: Search probabilities at mean values for mean and median daily commute costs (with other covariates)

	Mean dcc	R30	R40	R60	Median dcc	R20	R30	R40	R60
Pr(y=1x):	0.7013	0.6939	0.6862	0.6705	0.7014	0.6846	0.6659	0.6466	0.6065

The results reflect the possible negative impact of transport costs on search probabilities among unemployed individuals across South Africa. This was conducted over a range of indicators that could affect the search probabilities of South Africans. This study finds that as one moves from the bottom quarter to the top quarter of the transport cost distribution, search probabilities do not change much. Table 3.12 shows the 25th, 75th and 95th transport cost percentiles. It shows that, for both mean and median values, the probability of searching declines as transport costs rise. For instance, moving from the 25th to the 50th percentile results in a very small decrease in the search probability—0.48% for

mean values and 0.9% for median values. Moving from the 50th to the 75th percentile results in a search probability decrease of 1.1% and 1.94%, for mean and median values respectively. Although these are very small changes in the search probabilities of these individuals, it demonstrates the possible impact that transport costs may have on the search behaviour of unemployed South Africans.

Table 3-12:
Distribution of search probabilities at mean values for mean and median daily commute costs

	Mean dcc			Median dcc		
	25th	**50th**	**75th**	**25th**	**50th**	**75th**
	R4.34	R10.79	R25.44	R0	R5.2	R16
Pr(y=1x):	0.7132	0.7084	0.6974	0.7204	0.7114	0.692

3.8 Endogeneity and simultaneity

Several covariates are potentially endogenous. Kingdon and Knight (2006) suggest that education and location might be correlated with unobserved motivation, which in turn could affect search behaviour. Finding measures to control for this was difficult and was not done in this paper. The results could also suffer from selection bias, which may be caused by individuals self-selecting themselves into areas with higher employment rates or lower transport costs or both. This could subsequently exacerbate the potential simultaneity in the relationship between searching and transport costs. This is because the direction of causality is unknown, as the relationship can be positive or negative. This is because higher transport costs can account for increased searching, and increased searching creates higher transport costs. For instance, people may relocate to areas with higher employment, which drives down transport costs. This reflects a contrasting relationship to the one found in this study. This suggests that the results must be treated as correlations and, therefore, no causal claims about the relationship between searching and transport costs can be made here.

4 Chapter 4

4.1 Introduction

This chapter assesses whether there is a positive relationship between reservation wages and transport costs in South Africa. It seems likely that people will try and recoup some of the costs incurred when commuting to work. Thus, they require higher actual wages to take away the same wage, and this should be reflected in their reservation wages. Therefore, reservation wages may seem high, but only because individuals are factoring in these costs. Particular focus is placed on one aspect of transportation access, namely transportation costs. Various other costs are associated with transport that are not necessarily monetary, such as transport accessibility and safety and the length of the commute. These costs might affect the searching behaviour of many individuals, but many of these variables are not available in the dataset used and therefore cannot be included. This study only considers the pure monetary use-cost of transport used in searching for work. Understanding the effects that search costs may have on raising reservation wages, which in turn may raise actual wages, is important because this might help explain the high levels of unemployment in South Africa. If this is the case, then transport policy may be important in helping to create employment.

4.2 Transport costs and reservation wages

The reservation wage is the amount of income that makes a person indifferent to choosing between work and leisure (Killingsworth, 1983). It is also defined as the minimum acceptable wage for an individual to accept a job (Carolina and Pau, 2008). Studies, such as Killingsworth (1983), Carolina and Pau (2008) and Bhatti (2012), have shown that various indicators are correlated with the reservation wage. These indicators, which include numerous individual and household characteristics, can provide valuable insight into how individuals formulate their reservation wages. The classical participation model states that the reservation wage of the unemployed is a monetary value that compensates for leisure forgone by the individual (Mortensen, 1986). The data used in this paper, which was collected by the LMES, treats reservation wages similarly to what Carolina and Pau (2008) describe as the information-gathering approach to job search: it takes time and money to gather information about potential jobs. This is a costly exercise and assumes that an individual's reservation

wage includes these costs. The relationship between commute costs and wages stems from individual allocations of work or search and leisure time (Laird, 2006). This allocation is subject to time and budget constraints, and the resulting cost-benefit analysis centres around a minimum reservation wage that fits this budget. For individuals who live farther away from CBDs, the commute costs are higher and they must be compensated for this increased cost or face discouragement or unemployment. Laird (2006) conducts an in-depth analysis of the relationship between commute costs and the wage rate. He uses a two-stage least squares model and finds strong evidence for commuting cost compensation within the wage. The following section discusses studies that shed light on this relationship.

Although this paper primarily deals with transport costs and their effect on unemployment, it ultimately falls within the realm of spatial economics. Larger transports costs are attributed to longer commuting distances for job seekers and the employed. This distance is, by and large, a spatial issue that is affected by various factors, including agglomeration, urbanisation, infrastructure development (in terms of transport facilitation and utilities) and the general lay of the land. Earlier studies, such as Forster (1983), stress the relevance of the relationship between patterns of intra-urban unemployment and metropolitan spatial organisation. Forster examines the local unemployment rates in Adelaide (Australia) and argues that residential segregation and accessibility (transport and communication) are key contributors to unemployment. Both themes show how location affects someone's unemployment position. South Africa's past geographical segregation of people seems to be counter to the textbook model of urban residential choice, as described by Holzer (1991). Given the assumptions of a fixed central city and free choice, individuals or families choose an optimal residential location within a metropolitan area (Holzer, 1991). This choice mainly amounts to a decision between living closer to the city centre or in the suburbs, which are located farther away. Furthermore, this decision is largely dictated by individual tastes for better, and usually more expensive, central housing or longer commutes (Holzer, 1991). Those that choose longer commutes usually do so because of their desire for amenities associated with suburban lifestyles, such as, safety, pools, and gardens. This more relaxed lifestyle may be chosen by those who have, or are expecting, children, while their urban counterparts choose a faster paced lifestyle closer to the urban centre (Holzer, 1991). Holzer (1991)

asserts that, given the same skill level for suburban and urban individuals, there should be some compensation for the longer commutes experienced by those living farther away from their place of work. This compensation is indirectly achieved by lower rent costs for those living farther away, but this is counter balanced because they have to pay for longer commutes. This also implies that wages decrease as one moves from the city centre to peri-urban employment areas (Holzer, 1991). So, for a family who prefers peri-urban amenities and lifestyles, living and working away from the city centre would be ideal, since they also reap the benefits of shorter commutes. These choices were, and are still to some degree, non-existent for the previously segregated African populace of South Africa.

The freedom to choose their housing location, however, is a liberty that few possess (Holzer, 1991). Holzer refers to suburban zoning practices in America that were similar to those implemented in South Africa during Apartheid. These practices aimed to restrict low income families or racial groups from accessing certain areas. Constructing these locational divides meant that these people faced a greater burden, in terms of costs and employment opportunities. This is mainly because of the increased search costs and the decrease in the information about job opportunities (Holzer, 1991). The informal information networks that these individuals rely on to access potential jobs are less effective the farther away they are from these jobs (Holzer, 1991). Holzer asserts that the impact of this is even more severe for younger, less skilled job seekers who rely heavily on informal information networks to obtain employment. For those who do manage to attain employment, the increase in total commute costs lowers their net wages, in terms of both the monetary cost as well as the time spent working (Holzer, 1991). If these commute costs reduce net wages to below their minimum acceptable wage, then opting for continued search strategies or dropping out of the labour force become the only optiona available to them (Holzer, 1991).

Arnott (1998) asserts that two broad factors, commuting costs and job searching, are identified in the SMH literature and that these factors contribute to housing choices and the role that job suburbanisation has on employment outcomes. In the South African context, the SMH needs to be thought of in terms of its inverse: unskilled jobless people on the urban fringe seek jobs by traveling into suburban areas or city centres, and not outward from city centres. In either case, the commuting

costs are large because of the poor transport systems to and from these areas. This increases the monetary and time costs of commuting from these areas (Arnott, 1998). In terms of job searching, location has a large impact on attaining information relating to potential jobs (Arnott, 1998; Wasmer and Zenou, 2002). This effect is amplified if this information is mostly attained through social networks (word of mouth or acquaintances). Although treated as separate themes, job searching inevitably requires some form of transport, which costs time and money. For individuals forced to live in areas far away from employment, these costs can be too high to overcome.

To accommodate spatial searching, where individuals place themselves in locations that alter their probably of attaining work, Van Ommeren and Rietveld (2005) adjust their model to exclude random job arrival rates and replace this with search efforts. They assert that the closer an individual's residential location is to regions of employment, the greater the expected benefit, compared to those who are farther away. They suggest that this even implies that a negative relationship exists between job search effort and distance. This somewhat contradicts Van Ommeren (1998), who finds that footloose commuters are willing to increase their search efforts for longer commute times. This might suggest that job seekers presume that their efforts and increased commute times will be compensated by their wages. This could further indicate that an increase in commute costs and a decrease in housing service expenditure can be treated as substitutes, suggesting that the location decision does not affect the actual monetary cost of job searches (Rouwendal and Rietveld, 1994). The common view that longer commuting times (higher costs) are compensated by higher wages and lower housing prices is not shared by Van Ommeren (1998). He finds that workers are not deterred by small changes in commuting times and that they will reposition themselves in the labour market more rapidly than previous findings suggest. Zax (1991) produces different results. He estimates an earnings equation and finds that compensation for commuting comes from both the housing and labour markets through cheaper residential expenses from housing far from the CBD and increased wages for time spent commuting to the CBD. Zax (1991) also asserts that compensation for commuting costs is driven by residential mobility and individuals who lack residential mobility are compensated the least.

Arnott (1998) uses a trade theoretic model to formalise the SMH. He assumes, as does this paper, that transport costs are made up entirely of monetary costs. Although it is unrealistic to exclude other costs, such as time (or distance) or the fatigue caused by traveling, adding these differentiated costs would complicate things without contributing much (Arnott, 1998). Arnott also suggests that the time related to commuting costs can be approximately measured as a fixed proportion of the wage. Arnott (1998) asserts that travel time costs are more important than monetary costs, but chooses to omit this component of the commute cost. Ommeren and Fosgerau (2009) expand on this point. In their paper they suggest that many papers focus on the time component of commuting costs. However, they argue that although time is an important aspect of commute costs (contributing between 20% and 100% of the hourly gross wage with monetary costs comprising 30% to 40% of travel costs), the other components cannot be ignored. These include the mode of transport, deterioration of shoes, fatigue and increased food costs for those walking longer distances.

Small and Song (1992) resolve the discrepancy between Hamilton (1982) and White's (1988) papers. Essentially, Hamilton proposes that there is approximately a 90% excess of commute distances. This was later challenged by White, who only finds a 10% excess (Small and Song, 1992). A wasteful or excess commute is calculated by first measuring the average commute distance or time (in White's paper) and comparing this to the actual commute distance or time. Observations that are above average are considered wasteful. Small and Song (1992) conclude that Hamilton and White's use of time and distance in their data set had very little impact on the final results, and that the largest portion of disparity between their results is caused by the way in which the data was tested. Both White and Hamilton attempt to calculate excess commutes by using cost minimisation techniques, but Hamilton used the popular monocentric[3] framework to do so. Hamilton's monocentric model, which is CBD orientated (Richardson, 1988), relies on various assumptions, the most important of which is location equilibrium. This assumption implies that the marginal benefits and costs of moving away from the city centre are equal (Blackley and Follain, 1987). This is based primarily on the theory that by increasing the commute distance (and costs) between residence and job locations, one offsets these

[3] Model in which employment and population densities decline exponentially from a central point.

costs through decreased residential expenditure (Richardson, 1988). Having a single job centre and minimising costs based on the average commutes of individuals by using this model invariably produces overstated results, because there is a prevalence of intra-zone commutes (Small and Song, 1992). White (1988) aggregates the data using municipal zones. These municipalities cover a large area and they each have their own job centres. White (1988) uses similar minimisation techniques to calculate excess commutes. She uses these aggregated zones to allow the data to produce far fewer excess commutes because of the within zone commutes (Small and Song, 1992). Small and Song (1992) find that by disaggregating areas into smaller zones, they could find a much smaller discrepancy between the estimated and actual distance and time values. They managed to do this by comparing the estimates from small and large zones. Small and Song (1992) find that White's method generates understated commute times because of aggregation bias and the use of large zones in her data set.

Ommeren and Fosgerau (2009) state that commuting costs consist primarily of travel time and monetary costs. They use a dynamic search model, which aims to estimate the marginal cost of commuting by using the compensation workers receive for travel while in the labour market. Other components of commute costs may include those that affect individual satisfaction, like stress and the risk of accidents (Ommeren and Fosgerau, 2009).

Jun and Hur (2001) estimate the costs of commuting associated with the Seoul Greenbelt. The Seoul Greenbelt is a forestry belt created in four phases as part of the South Korean development plan implemented from 1971 to 1982. This greenbelt created satellite towns and significant effects on commuting patterns and costs. To estimate these costs, Jun and Hur (2001) create a scenario where the greenbelt does not exist and reallocate the jobs and workers (from either greenbelt boundary) among zones within the greenbelt that would theoretically have developed. They then calculate job and residential densities and apply them to this scenario. Jun and Hur (2001) use the Fractar method (a type of trip distribution model) to estimate commute costs. Two main components of this estimated commute cost consist of the amount paid for transport (out-of-pocket expenses) and the total travel time. They then apply these numbers to work done by Lee (1988), who calculates the out-of-pocket

expenses and total commute costs split among various modes of transport. They find that individuals living or working within the greenbelt areas incurred significant increases in commute distances and costs. Although Jun and Hur (2001) do not directly address the objectives of this paper, their paper serves as an example of how commuting costs for people in decentralised areas, such as those in rural areas, are far larger. The greenbelt example is similar to the experience of rural dwellers in South Africa, since they have to traverse areas with poor road infrastructure that are far from CBDs.

Rupert et al. (2009) argue that commute time and distance affect reservation wage determination and the decision to accept or reject job offers. The notion that there is a spatial component, and resulting commuting costs, to the job search decision has been well documented (Rupert et al., 2009). However, macroeconomic theory does not include many studies that apply commute costs to search theory (Rupert et al., 2009). Rupert et al. develop a job search theory where jobs consist of a productivity and commute distance component. They argue that a positive relationship exists between commute distance and wages. Furthermore, a negative relationship exists between productivity and commute distance, because productivity relies heavily on the vigour of workers and increased commuting distances invariably tire workers (Rupert et al., 2009). However, it is possible that high paying jobs located in remote areas could be lucrative enough that individuals are willing to overcome the increased commute distance (Rupert et al., 2009) and transport costs. A problem that they acknowledge their paper faces is that it uses a data set with no reservation wages. They solve this problem by using data on accepted wages together with implied commuting costs, which are conditional on employment, to infer a relationship between commuting distance and reservation wages. By doing so, they estimate a reservation wage strategy. Using a conditional (recursive) mixed process approach,[4] they suggest that a large and positive relationship exists between commuting time and wages. Although the relationship increases at a decreasing rate, they find that for every hour of commuting, individuals need to be compensated by a 28.5% increase in their wages. However, standard OLS estimates show a much smaller required commute time to wage compensation of 8.4% to 9%. Rupert et al. (2009) assert that this outcome implies that commute time affects the bargaining

[4] "…a simultaneous equation system where the different equations can have different kinds of dependent variables" (Rupert et al., 2009, p.9)

power of workers. They propose that higher wage compensation, in the form of transport costs, is received by workers with lower bargaining power. This suggests that workers with lower skills, who exist in abundance, are more likely to be compensated for their commuting costs than their highly skilled counterparts.

4.3 Data—The Labour Market Entry Survey

This chapter uses the LMES data set instead of the NIDS, since this data set has better information for analysing reservation wages and commute costs. The LMES was collected as part of a randomised control study aimed at evaluating the impact of a wage subsidy on employment outcomes of randomly selected participants. The study, which spanned four years, included three provinces in South Africa: Kwazulu-Natal, Limpopo and Gauteng. A total of 4000 respondents, aged between 21 and 25 years, were interviewed in 2009, of which 2500 were identified in selected enumeration areas. Figure 4.1 gives a cursory impression of the distribution of observations when the three provinces were first sampled in 2009.

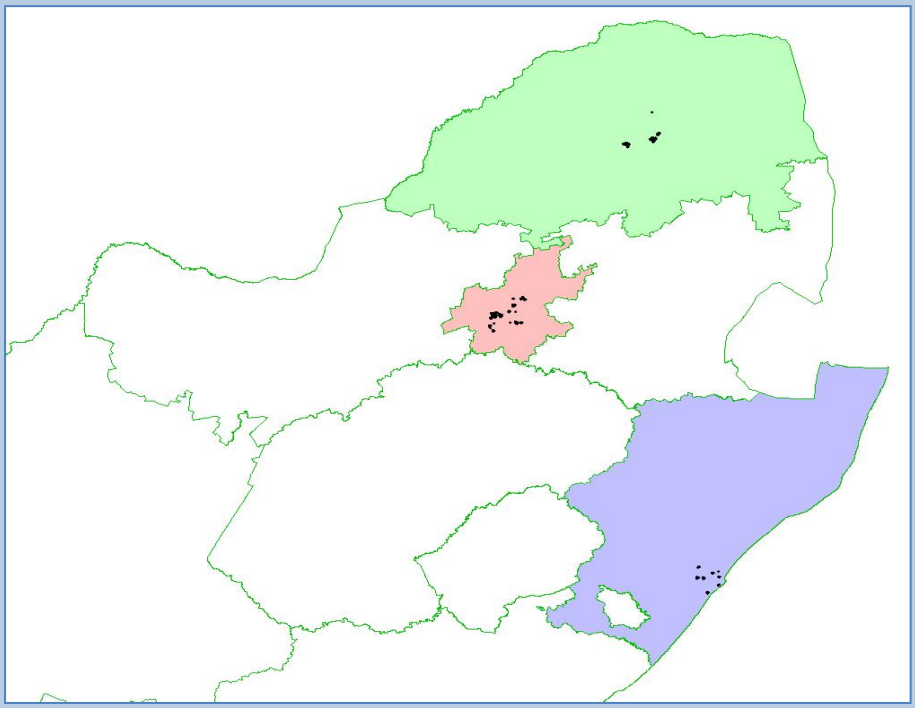


Figure 4-1: Observation distribution of the LMES (2009)

More than half of the sample was identified in Gauteng, with the remaining respondents being almost evenly distributed between Kwazulu-Natal and Limpopo. The remaining 1500 respondents came from

Department of Labour labour centres located near the enumeration areas. The survey ended in 2012, but at the time of writing only the 2011 data set was complete, which includes 1367 observations from that year. This data is divided into a treatment and control group, with almost 50% of respondents in each group. Allocation to the two groups was independent of employment status, and a similar number of employed, unemployed and discouraged workers were distributed between the groups. Members of the treatment group were given wage subsidy vouchers to encourage them to look for employment. Members of the control group were left to their own devices. This may distort the results because their transport costs would likely have been lower had they not received the voucher. Table 4.1 shows a summary of the transport costs by treatment group. It shows that the influence of the treatment group is not a concern. This is because much of the variation caused by transport costs occurs within each group (treatment or control) and not between each group.

Table 4-1: Median transport costs by treatment group

Variable		Mean	Std. Dev.	Min	Max	Observations
Transport costs	overall	428.777	168.4558	150	1100	N = 971
	between		10.78754	436.3971	7700	n = 2
	within		168.2828	142.3799	1107.636	T-bar = 485.5

The LMES currently comprises four years of data. However, certain questions that are relevant to this paper were only included in 2010, and are still part of the questionnaire. This study uses a cross sectional analysis of this data. This paper uses the latest 2011 LMES data, for which there are 2358 observations available. The sample included labour centre respondents. These individuals were removed from the sample because they were not assigned to a particular enumeration area and many of them came from provinces other than those originally sampled in the LMES survey. Practically, this meant that it was impossible to create cluster wide transport cost estimates for these individuals, which is a key variable of interest in this analysis. Excluding these individuals reduces the number of observations to 1367. Individuals who had moved out of the original enumeration area were also excluded. This was done for a similar practical reason—cluster wide transport costs could not be calculated for these individuals because these they reported transport costs for areas outside of the

primary sampling clusters. This also mitigates possible selection bias within the sample that may exist from respondents selecting themselves into areas where transport costs may be lower and success in the labour market more likely. Removing these observations created a final sample with 1092 observations. Table 4.2 shows the distribution of observations among provinces by gender. Of the respondents, 60% are women and 40% are men. This is quite consistent for each of the three provinces. Kwazulu-Natal was a difficult province in which to gather and retain observations. This can be seen in the disproportionate reduction in numbers. It should represent around 25% of the sample, but only represents 11.7%.

Table 4-2: LMES—Gender distribution by province

	Gender								
Province (psu)	Female			Male			Total		
	No.	Col%	Row%	No.	Col%	Row%	No.	Col%	Row%
Gauteng	367	57.2	57.1	276	61.6	42.9	643	59	100
Kwazulu-Natal	**78**	12.1	60.9	**50**	11.2	39.1	128	**11.7**	100
Limpopo	197	30.7	61.8	122	27.2	38.2	319	29.3	100
Total	642	100	**58.9**	448	100	**41.1**	1090	100	100

The LMES data can be divided into unemployed and employed groups. This distinction is evident from two main questions and numerous control questions throughout the LMES. However, this analysis only considers two questions. The first question (Question 1) asked, *"What currently takes up most of your time?"* Six options were given, of which four, indicating employment or lack thereof, were used. Only individuals who were employed or unemployed were considered, reflected by a 1 or 0, respectively. The second question (Question 2) was asked later as a control. It is considered more of an official definition and asked *"In the last week did you work for a wage, salary, commission or any payment in kind, (including paid domestic work), even if it was for only one hour?"* There were only two responses to this question: yes or no. Table 4.3 summarises the responses to the two questions. Both questions reflect a 40% to 60% split between each state. There is some overlap between these questions and both are used separately to see if this significantly affects the final outcomes.

Table 4-3: Reported employment status

Question 1	Question 2			
	Unemployed	Employed	.	Total
Unemployed	714	104	156	974
Employed	29	355	9	393
Total	743	459	165	1,367

Each group was asked questions pertaining to a wide range of household, personal and geographical indicators. In some instances unemployed individuals were asked questions that the unemployed were not and vice versa. Questions pertaining explicitly to transport costs in the LMES were only posed to employed respondents. This means that if the entire sample is used, transport costs must be generated for the unemployed individuals. A cluster wide average transport cost was created to overcome this problem. Respondents were assigned a mean transport cost from each cluster based on the reported transport values.

4.4 Methodology

This chapter uses an OLS regression analysis to explore the effects that transport costs have on reservation wages. More specifically, it investigates whether transport costs are associated with an individual's reported reservation wage. A priori, it is expected that the latter increases with the former. This study hypothesises that transport costs do not increase reservation wages. The OLS regression analysis is similar to a Mincerian-based earnings equation that uses commute costs as an independent variable.

4.5 Descriptive analysis (econometric specification)

This section describes the relationship between reservation wages and transport costs by using a standard OLS regression analysis with reservation wages as the dependent variable and transport costs and other individual characteristics as independent variables. This method is very similar to a Mincerian earnings equation. However, the dependent variable is the log of reservation wages and transport costs are introduced into the model. In Mincer (1974), education and experience form the central components of the Mincerian wage function. Bhatti (2012) asserts that social, demographic,

regional and economic factors are now commonly added to this function. Table 4.4 summarises the variables used in this analysis.

Table 4-4: Summary of dependent and independent variables

Variable	Label	Definition
Dependent variables		
Reservation wage (far)	*lnrwfar*	*Log of the reported minimum wage*
Reservation wage (near)	*lnrwnear*	*Log of the reported minimum wage for jobs nearby*
Independent variables		
Transport costs	*median_trans*	*Cluster wide median transport cost*
Education		Dummy variable for education with *edu1* as the base
	edu1	*Only schooling*
	edu2	*Certificate/Diploma with LESS than grade 12*
	edu3	*Certificate/Diploma with grade 12*
	edu4	*Bachelor's degree or higher*
Matric		Dummy variable for matric with *nomatric* as the base
	nomatric	*No matriculation as a minimum*
	matric_ne	*Matriculation as a minimum but no exemption*
	matric_e	*Matriculation as a minimum with exemption*
Work experience		Dummy variable for full time work experience with *workexp1* as the base
	workexp1	*Between no work and 3 months of work*
	workexp2	*Between 3 months and 1 year*
	workexp3	*Between 1 and 3 years*
Part time experience		Dummy variable for part time work experience with *ptworkexp1* as the base
	ptworkexp1	*Between no work and 3 months of work*
	ptworkexp2	*Between 3 months and 1 year*
	ptworkexp3	*Between 1 and 3 years*
Gender	*Male*	*Gender variable with female as the base*
Rural	*rural_urban*	*Respondent resides in a rural area*
Informal settlement	*informal*	*Respondent resides in an informal settlement*
Age	*age*	*Respondent's age*

4.5.1 Dependent variables

Three LMES questions capture the minimum wage. It is possible to use this data to indirectly capture the premium that people place on job searching, which we assume includes mostly transportation costs. In areas such as Limpopo, KwaZulu-Natal and some parts of Gauteng, the probability that

individuals are unemployed is high. Coupled with fewer skills, as well as short employment durations, these individuals are unlikely to move easily between employment states. Therefore a static analysis is used to measure the current value. The following sections describe the two methods that are used to explore this relationship, and describe the included variables and the reasons why they are included. Two different reservation wages were reported in the LMES. The first measured the minimum wage for jobs and the second measured the minimum wage for jobs nearby.

- *What is the MINIMUM MONTHLY wage you are prepared to work for 8 hours a day, 5 days a week for jobs NEAR to your home?*
- *What is the MINIMUM MONTHLY wage you are prepared to work 8 hours a day, 5 days a week for?*

Both are included to determine whether the results are consistent. The natural logarithm of reservation wages is calculated to minimise the effect of outliers on regression estimates. The log[5] of each variable is calculated to reduce the impact of these outliers. The results in Table 4.6 show these changes. Most of the variation still comes from within the clusters, but to a much smaller degree.

Other possible forms of the dependent variable were examined before deciding which to use in the final analysis. Another variable in the LMES describes the minimum wage that respondents were willing to accept if they were desperate for a job. The LMES asked "*What is the MINIMUM MONTHLY wage you are prepared to work 8 hours a day, 5 days a week for if you were desperate for work?*" This question was included to test each respondent's understanding of the previous reservation questions posed to them, but it is likely that some accuracy of the question was lost because it was asked together with other similar questions. Walker (2003) notes that respondents are likely to report a fair wage instead of their minimum wage, since respondents enter a bargaining state when faced with those questions. Table 4.5a shows that a sharp drop occurs in the minimum wage, suggesting that the previous questions are not accurate indicators of the reservation wage. The three reservation wage measures were adjusted by calculating the log of these variables. The standard deviation of *rwfar* and *rwnear* are both around R2800, which is far greater than the variation in

[5] Each time a logarithm of any variable was created, the number one was added to the variable. For example, var1 = log (var +1). This simple linear transformation technique is used to avoid trying to log zeros, which is impossible, and has no significant effect on the variable.

rwdesp. Natural logarithms of each reservation wage are used. The desperate wage in the LMES is theoretically a far more convincing indicator of the reservation wage. However, this variable must be used cautiously because, in their desperation, respondents might have unwittingly ignored commute costs and only considered those costs that they require to service their basic needs.

Table 4-5a: Summary statistics for indicators of reservation wages

Stats	*rwfar*	*rwnear*	*rwdesp*
Mean	4241.473	3231.08	1904.979
Median	3500	2500	1500
Stv.Dev	2818.717	2832.745	1118.323
Min	500	500	50
Max	25000	60000	8000
Obs.	1356	1359	1301

Table 4.5b compares the summary statistics of the unemployed and discouraged responses to the *rwfar*, *rwnear* and *redesp* variables. This is important because of the significance placed on reservation wages and the potential differences that exist between unemployed and discouraged individuals. For all types of reservations wages, the unemployed group reported higher mean wages. For *rwnear*, this is around 54% larger for the unemployed group. The higher unemployment figures suggest that this group is more confident about their ability to find work that suits their experience or education levels.

Table 4-5b: Reservation wage differences by employment status

	rwfar		rwnear		rwdesp	
Stats	Discouraged	Unemployed	Discouraged	Unemployed	Discouraged	Unemployed
Mean	3847.82	5227.132	2830.358	4360.87	1798.331	2172.318
Median	3000	4000	2500	3500	1500	2000
Stv.Dev	2392.8	3488.428	1897.125	4332.445	1055.071	1224.311
Min	500	1000	500	800	50	480
Max	16000	25000	16000	60000	8000	8000
Obs.	969	387	973	391	930	371

4.5.2 Independent variables

The LMES captured transport costs with a single question: "*How much does transport cost per month?*" However, this question was only answered by employed individuals, of whom there were

only 409 at the start of the study. Removing individuals who had moved from locations where they were sampled left only 318 observations in 2011.

To make the transport costs comparable, individuals with similar employment durations and those who worked full-time (since these types of people are most likely to give an indication of monthly transport costs) were selected.[6] Two questions were used to identify these people: *"What is the average number of days worked a week at [place of work]?"* and *"What is the average number of weeks worked in a month at [place of work]?"* Isolating individuals who worked five days a week for four weeks was done by using both questions. Table 4.6 shows a summary of the adjusted transport costs' construction distinguished by gender. Of the 318 respondents, only 139 worked five days a week for four weeks. Attempts to create a variable that includes all 318 observations failed to yield any significant results. This is probably due to unreasonable responses given by respondents for various employments states that were eliminated in the final variable.[7] The average cluster-wide transport cost—*median_trans*—essentially matches the transport costs of the employed and unemployed by cluster, therefore avoiding excluding the unemployed from the final estimations. Table 4.6b shows *median_trans* for the entire sample after construction by gender.

Table 4-6a: Monthly transport costs by gender

Gender	Female	Male	Total	Female	Male	Total	Difference
Obs.	156	162	318	61	78	139	179
Median	380	400	400	400	437.5	400	0
Mean	419.2051	526.716	473.9748	461.5246	648.7179	566.5683	-92.6
Std.Dev	292.771	729.7133	561.4703	283.1412	982.3664	762.8987	-201.4
Min	5	5	5	100	50	50	-45
Max	1800	8000	8000	1800	8000	8000	0

Table 4-6b: Median transport costs by gender

Gender	Mean	Median	Stv.Dev	Min	Max	Obs.
Female	403.1939	400	139.0171	88	800	923
Male	418.8172	420	140.787	88	800	640
Total	409.5912	400	139.9108	88	800	1563

[6] The unadjusted transport costs were used in trial estimations, the results of which were not very different from the adjusted transport costs.
[7] Using the un-adjusted figures only produced significant results for the employed respondents who originally answered the question. Creating a mean cluster wide un-adjusted figure had no significant results.

Figure 4.2 shows the density of median transport costs. Most observations are distributed around R400 per month. The median transport cost is approximately R410 per month, which is commensurate with what is expected of travels costs to and from these work locations. 119 observations across seven clusters were dropped. All the densities for the independent variables compared to the median transport costs generally follow the same type of distribution shown in Figure 4.2. This is not surprising, because these transport cost measures are all at the cluster level. This supports the use of this measure of transport cost because of the small variation across clusters.

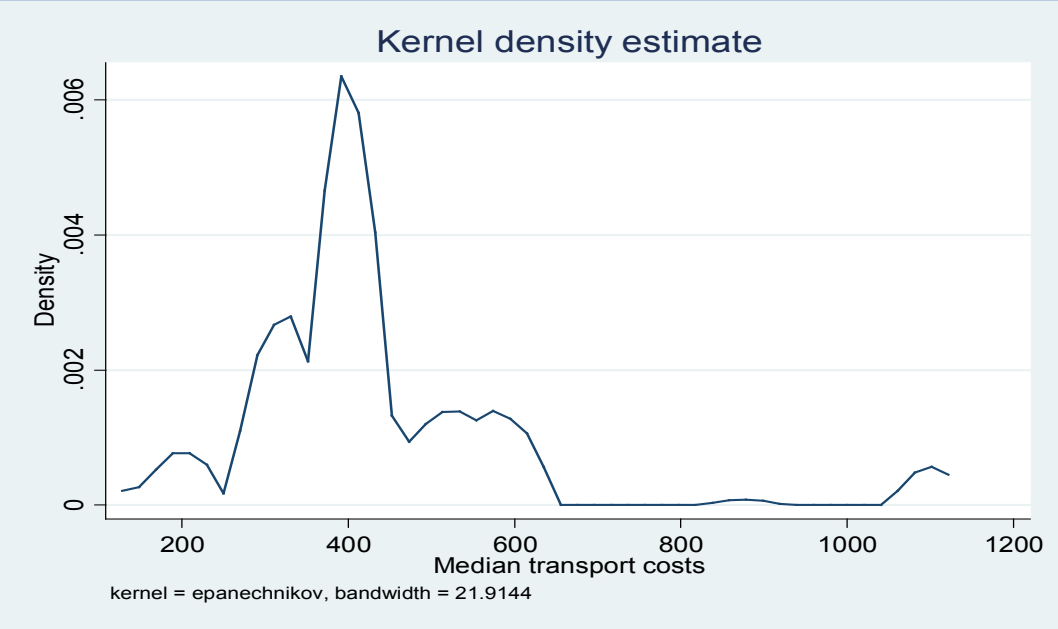


Figure 4-2: Density of median transport costs

Figure 4.3 is a symplot of median transport costs. It shows the distance of observations above and below the median. It is a way to identify outliers, which are observations that are not clustered around the median line. The data includes notable outliers, but removing them would harm the proceeding analysis because it will make the data less representative of the original sample.

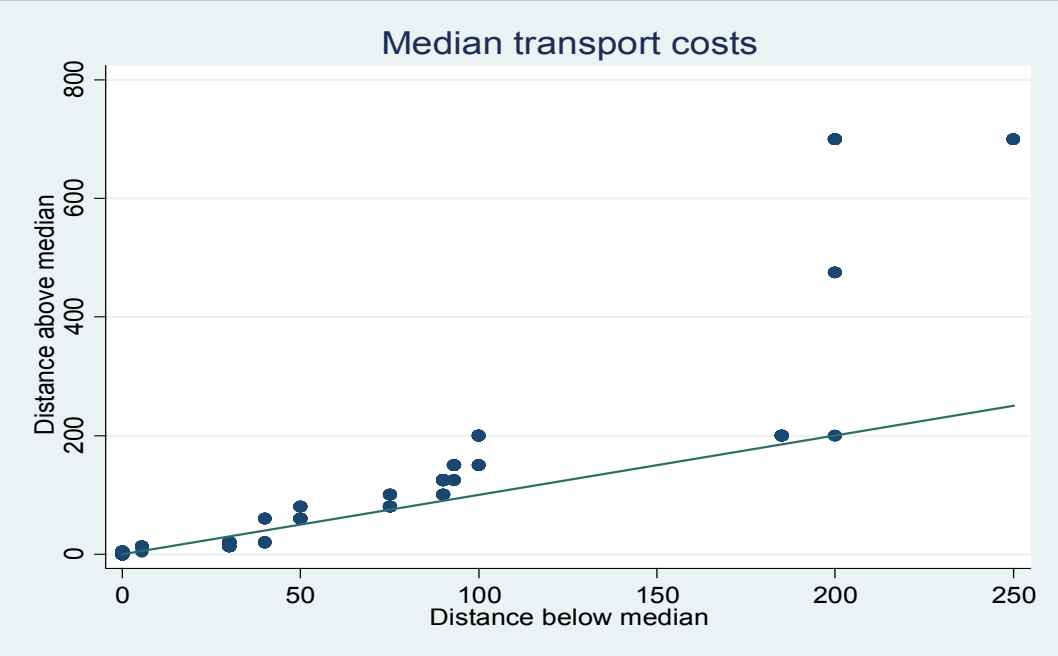


Figure 4-3: Symplot of median transport costs

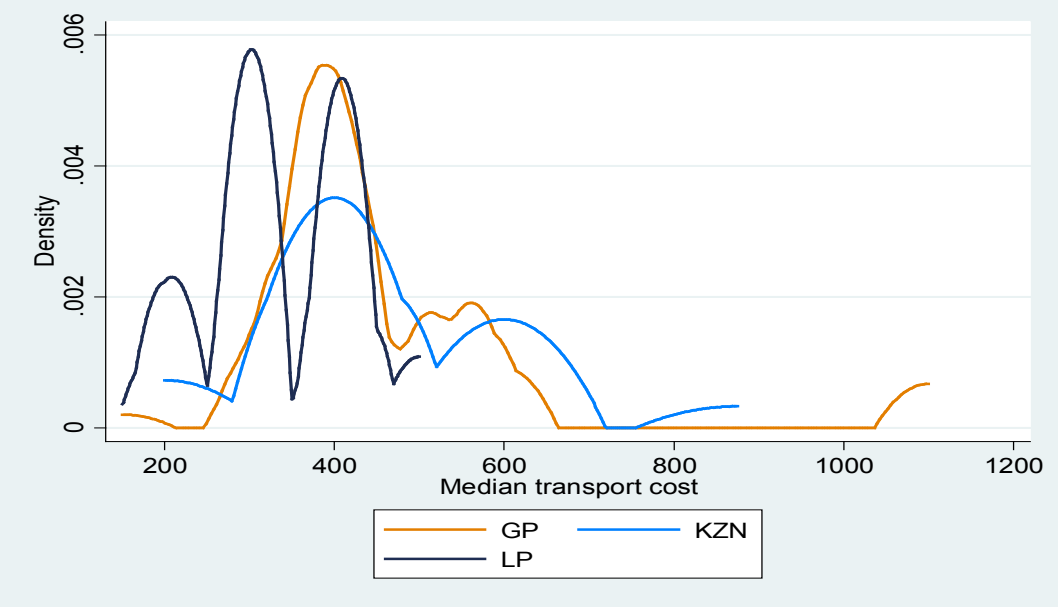


Figure 4-4: Median transport costs by province

Mincer's 1974 logarithmic earnings model implies that reservation wages should follow a linear function for years of education and a quadratic function for potential experience (Hou et al., 2011). However, the relationship between education and reservation wages is probably non-linear because of education's endogenous nature (Hou et al., 2011). Furthermore, given the heterogeneous nature of schooling years, it is unwise to emphasise the returns to education (Hou et al., 2011). The argument

for returns to education is double sided. It is generally accepted that the more education one attains, the higher the wage premium one expects. This means that an individual can only maximise their life-time utility by attaining a job that is commensurate with their human capital and monetary investment (Hou et al., 2011). If an individual cannot obtain commensurate employment, then they face a sunk human capital and monetary investment cost. Using dummy variables for categories of education allows for non-linearities. Four separate education variables are included in this model and were created from the question, *"What is the highest level of education that you have successfully completed?"* Individuals could choose from one of the five options, which separated respondents by those who had completed only primary schooling; a certificate or diploma with less than Grade 12; a Grade 12 certificate; a Grade 12 diploma; or a bachelor's degree or higher. The five categories were collapsed into four education categories by combining individuals with Grade 12 certificates or diplomas into a single variable. The four variables are called *edu1*, *edu2*, *edu3* and *edu4*. Figure 4.5 shows that a large portion of individuals, regardless of their education status, have a reservation wage between R1500 and R6000. The smallest group are those with bachelors' degrees or higher, who generally reported reservation wages between R1500 and R15000.

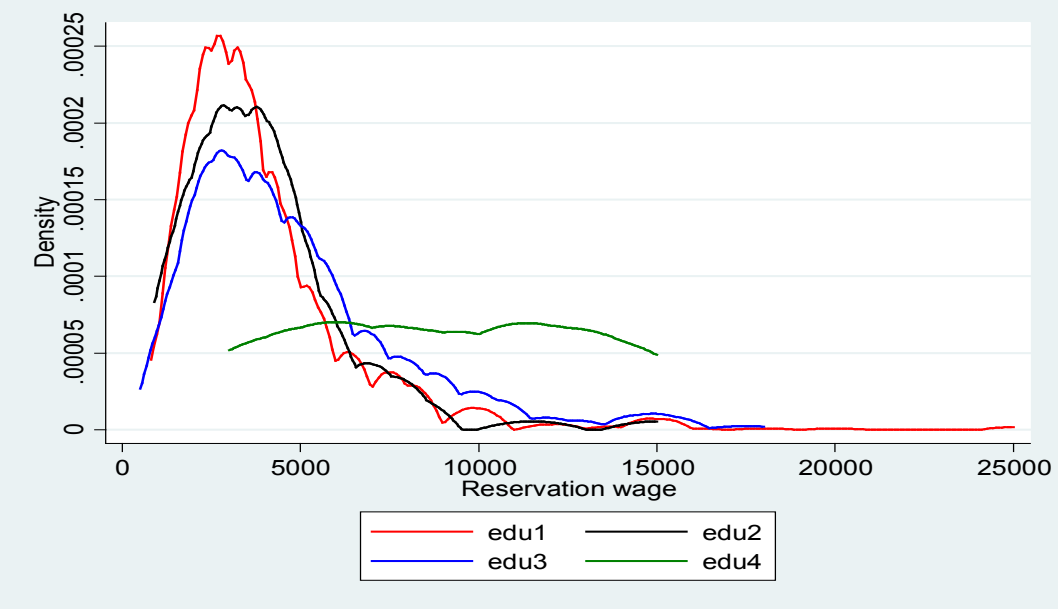


*Figure 4-5: Reservation wage (*rwfar*) by education level*

Another education variable which is thought to be a good indicator of job attainment is a matric certificate. In South Africa, a matric certificate is considered the basic requirement that individuals need to possess to be considered for employment. Matric might signal to potential employers that an individual has undergone a standardised test, allowing all candidates to be compared. Knowing this, individuals may find it more worthwhile to seek employment and therefore spend more time and money (transport costs) trying to attain a job. This status is captured by the question, *"Do you have a Matric certificate?"* If respondents answered yes, the survey asked whether they attained exemption into university. All three categories are identified by the variables *nomatric* (no matric), *matric_ne* (matric with no exemption), *matric_e* (matric with exemption). Table 4.7 shows that there are more females with matric. However, university exemption is equally distributed between genders.

Table 4-7: Matric certification by gender

	Do you have a Matric certificate?			
Gender	No	Yes - without university exemption (endorsement)	Yes - with university exemption (endorsement)	Total
	No.	No.	No.	No.
Female	315	392	100	807
Male	197	279	84	560
Total	512	671	184	1,367.0

Two categories of productive characteristics were captured by two questions relating to work experience. Kingdon and Knight (1999) find that experience increases with wages at a decreasing rate. The outcome here is expected to be the same, regardless of the type of experience that the individual has. Two main categories of work experience are created to measure experience: full and part-time work experience. These were captured by the following questions: *"How much work experience in TOTAL do you have working FULL-TIME for PAY at a person/business/government/non-profit/family firm?"* and *"How much work experience in TOTAL do you have working PART-TIME for PAY at a person/business/government/non-profit/family firm?"* Both questions had eight responses: less than 1 month, between 1 and 3 months, between 3 and 6 months, between 6 months and 1 year, between 1 and 2 years, between 2 and 3 years and more than 3 years. These responses were used to create three categories: 0 to 3 months' work experience, 3

months to 1 years' work experience and 1 to 3 years' work experience. Figures 4.6 and 4.7 show the distribution of observations for full-time and part-time work experience for reservation wages, respectively. Most individuals, regardless of their work experience, have reservation wages between R4000 and R5000. This is also true for both types of work experience. This is possibly because some individuals have both types of work experience. A large proportion of individuals clearly fall into the 0 to 3 months' work experience category.

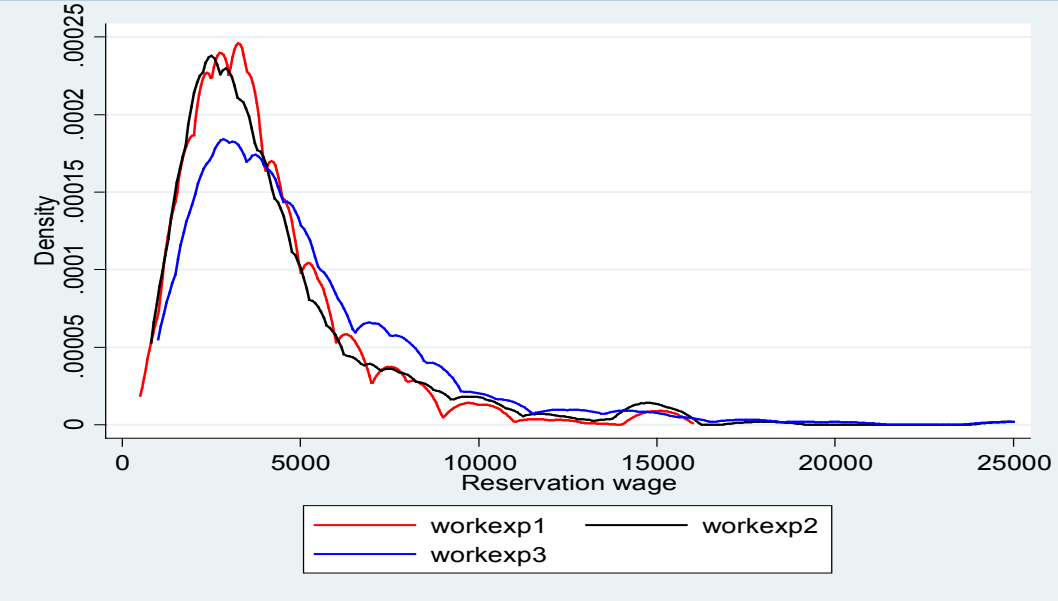


*Figure 4-6: Reservation wage (*rwfar*) by full-time work experience*

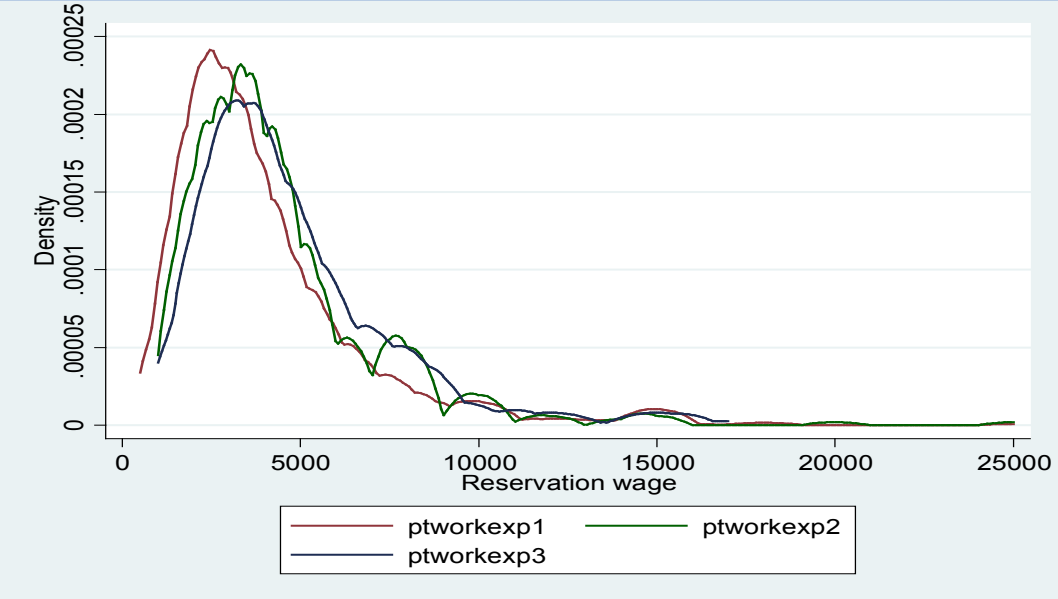


*Figure 4-7: Reservation wage (*rwfar*) by part-time work experience*

Age can be an important determinant of reservation wages, because older individuals are more likely to spend to spend more money on commuting to search for work or commuting for job requirements. It is plausible that an older unemployed individual could place more pressure, inadvertently or not, on the household to assist in the search for work because they have reached an age where contribution to the household has become a necessity. Young people are more likely to get away with loafing than older individuals. The LMES was presented to people between the ages of 20 and 24 years. Although this is probably not enough of an age gap to find something meaningful, it was possible to construct a smoother, more continuous version of the age variable by using the date of birth (day, month and year). Figure 4.8 shows this. The smoother red line represents the newly constructed age variable. An age squared variable was also constructed and included in the analysis because, a priori, there is a turning point in a person's age (around middle age) when the household feels that the individual should take care of themselves by acquiring a job. Whether age and age squared affect the transport premium in the same way as they affects wages is uncertain, and not much theory on its influence on transport premiums exists.

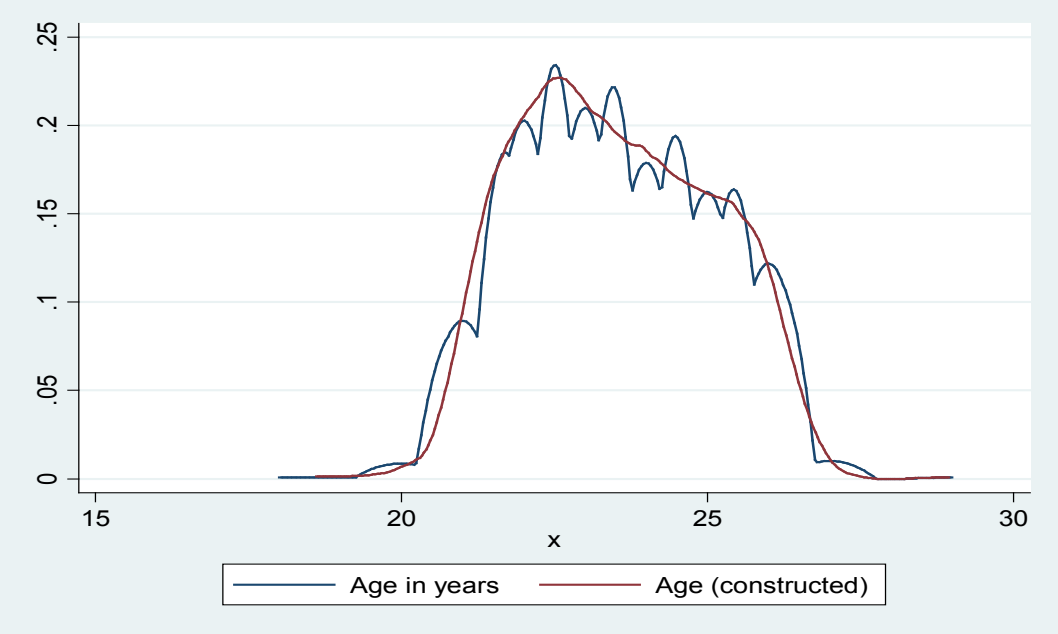


Figure 4-8: Distribution of age

Based on Kingdon and Knight (2006), some remoteness variables were included. These variables are ideal since they reduce endogeneity. They are not strictly exogenous because people may still choose

where to live. These variables help indicate whether distances from central employment areas hinder an individual's willingness to commute or search for work. Three dummy variables are included. They identify whether individuals come from rural or urban areas (*rural)*, whether they reside in an informal settlement (*informal*) and their sampling province. Both *rural* and *informal* were included in the data post collection specifically for this analysis[8] (see the appendix for a complete breakdown). The individual's *sampling province* was resourced for location variables associated with the survey and ascertains if certain provinces are better off than others in terms of transport cost allocation.

4.6 Results and analysis

This analysis focuses on the relationship between reservations wages and median transport costs. Table 4.8 shows the first of three separate regressions for the far (*lnrwfar*), and near (*lnrwnear*) reservation wages Three groups of regressions were run on the portion of the sample that was considered employed or not. The QLFS versions were used since they provide a more accurate and formal definition of employment status. For completeness, regressions were compiled that accommodated the entire sample based on the LMES definitions of employment (see Table A4.1 in the appendix). Table 4.9 shows separate results for the employed and unemployed to isolate which group is driving the results. The independent variables included in these regressions are: median transport costs (*median_trans*), gender (male), age, matric exemption, highest education level (*edu*), full-time work experience (*workexp*) and part-time work experience (*ptworkexp*). The regression output in Table 4.8 shows that *median_trans* for *lnrwfar* is significant at the 5% level. *median_trans* in both regressions is positive, which confirms that transport costs are associated with reservation wages. Since these are log-lin regressions, this output suggests that individual reservation wages increase by about 0.0228% for every rand per month spent commuting to jobs that are faraway. For example, a R100 increase in transport costs is associated with a 2.28% higher reservation wage. Interpreting coefficients with an untransformed dependent variable requires using the anti-log of the constant. In the first column, this amounts to approximately R1550 for all base categories in the regression. In other words, R1550 is the mean reservation value for those who are female, have no

[8] The table was constructed with the help of Hlayisani Rikhotso, who travelled to all three provinces conducting surveys for the LMES.

matric, no schooling, 0 to 3 months' full-time work experience, 0 to 3 months' part-time work experience, reside in an urban area and were sampled from Gauteng province. These results suggest that for every Rand that an individual spends on travel, they expect 35 cents in return. Put differently, individuals expect a 35% return on every rand they spend traveling to job locations. This suggests that individuals who commute on a monthly basis are not trying to recoup all their transport costs. Although *median_trans* is not significant at the 10% level, the coefficient estimate indicates that individuals expect a return of around 18%, which is approximately half of the return expected for jobs that are far. The relationship between transport costs and reservation wages may be smaller for closer jobs, since people may be able to walk to these types of jobs, or they may actually be factoring in the additional costs, including time and effort, which may be associated with commuting farther.

Positive relationships exist for gender, age and all education related variables, which is expected. However, only the coefficient for individuals with a bachelors' degree or higher is significant for both regressions. Being in a rural area is only significant for reservation wages for jobs that are close and decreases the reservation wage. This suggests that these individuals value themselves less because they are more likely to have fewer skills than their urban counterparts. Individuals living closer to the CBD might have access to more information about multiple jobs, which gives them a wage negotiation or job choice advantage. The only province of birth that is significant is Limpopo, and it has a negative relationship with reservation wages. This is not unusual because many individuals from Limpopo seek employment in Gauteng, and they may feel that they have to reduce their asking rate to compete with others. However, this should be true for almost all provinces (regardless of significance) relative to the Gauteng province (base category), therefore there may be other psychological or sociological factors at play.

Table 4-8: OLS estimation for the employed and unemployed

VARIABLES	lnrwfar coef	se	lnrwnear coef	se
Median transport costs	0.000228**	(0.000112)	0.000141	(0.000117)
Gender: male = 1	0.146***	(0.0366)	0.195***	(0.0384)
Age (constructed)	0.0192	(0.0117)	0.0170	(0.0123)
Matriculation as a minimum with exemption	0.439***	(0.0651)	0.449***	(0.0682)
Matriculation as a minimum without exemption	0.271***	(0.0437)	0.249***	(0.0460)
Certificate/Diploma with less than grd12	0.103	(0.0775)	0.0160	(0.0818)
Certificate/Diploma with grd12	0.0278	(0.0436)	0.0165	(0.0457)
Bachelor's degree or higher	0.574**	(0.223)	0.535**	(0.219)
Between 3 months and 1 year full-time work	0.0361	(0.0446)	0.0458	(0.0470)
Between 1 and 3 years full-time work	0.137***	(0.0466)	0.183***	(0.0488)
Between 3 months and 1 year part-time work	0.103**	(0.0409)	0.0988**	(0.0430)
Between 1 and 3 years part-time work	0.152***	(0.0541)	0.191***	(0.0566)
Rural_urban	-0.117	(0.0851)	-0.222**	(0.0896)
Informal settlement	-0.0302	(0.0392)	-0.0162	(0.0412)
Kwazulu-Natal	-0.00306	(0.0699)	-0.00112	(0.0732)
Limpopo	-0.120**	(0.0487)	-0.147***	(0.0512)
Constant	7.346***	(0.281)	7.131***	(0.295)
Observations	902		909	
R-squared	0.157		0.170	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

The estimations in Table 4.9 try to determine which portion of the sample is driving the results by distinguishing between the employed and unemployed groups. If the employed are including their transport costs in their reported reservation wages and the unemployed are not, then there will be a correlation between transport costs and employment. This shows that the results are larger and more significant for the employed part of the sample. This may be because the employed actually incur the transport costs (from which the variable is calculated). However, *median transport costs* for the unemployed for jobs that are far is also significant at the 10% level.

Table 4-9: OLS estimation for separate groups

VARIABLES	Employed lnrwfar	se	Employed lnrwnear	se	Unemployed lnrwfar	se	Unemployed lnrwnear	se
Median transport costs	0.000596**	(0.000254)	0.000814***	(0.000279)	0.000207*	(0.000124)	6.49e-05	(0.000127)
Gender: male = 1	0.156**	(0.0656)	0.193***	(0.0722)	0.136***	(0.0446)	0.182***	(0.0454)
Age (constructed)	-0.00489	(0.0217)	0.00405	(0.0239)	0.0259*	(0.0138)	0.0226	(0.0141)
Matriculation with exemption	0.499***	(0.113)	0.463***	(0.124)	0.380***	(0.0796)	0.396***	(0.0810)
Matriculation without exemption	0.373***	(0.0862)	0.293***	(0.0954)	0.236***	(0.0504)	0.222***	(0.0516)
Certificate/Diploma with less than grd12	0.168	(0.177)	0.0611	(0.196)	0.0939	(0.0856)	0.0151	(0.0876)
Certificate/Diploma with grd12	0.185**	(0.0745)	0.173**	(0.0820)	-0.0552	(0.0527)	-0.0627	(0.0537)
Bachelor's degree or higher	0.582*	(0.307)	0.400	(0.340)	0.480	(0.314)	0.579**	(0.281)
Between 3 months and 1 year full-time work	0.00790	(0.0821)	0.0483	(0.0908)	-0.0126	(0.0536)	-0.0205	(0.0547)
Between 1 and 3 years full-time work	-0.0132	(0.0791)	0.0940	(0.0872)	0.110*	(0.0618)	0.0945	(0.0628)
Between 3 months and 1 year part-time work	0.0585	(0.0730)	0.0866	(0.0805)	0.0687	(0.0497)	0.0428	(0.0508)
Between 1 and 3 years part-time work	0.0454	(0.0917)	0.200**	(0.100)	0.170**	(0.0672)	0.141**	(0.0685)
Rural_urban	0.110	(0.153)	0.0778	(0.169)	-0.218**	(0.101)	-0.355***	(0.104)
Informal settlement	-0.0508	(0.0743)	-0.0317	(0.0820)	-0.0195	(0.0457)	-0.000403	(0.0467)
Kwazulu-Natal	-0.241*	(0.133)	-0.188	(0.145)	0.0780	(0.0815)	0.0797	(0.0834)
Limpopo	-0.0425	(0.0957)	-0.0192	(0.106)	-0.122**	(0.0557)	-0.163***	(0.0568)
Constant	7.852***	(0.532)	7.253***	(0.586)	7.212***	(0.330)	7.053***	(0.335)
Observations	257		260		645		649	
R-squared	0.234		0.210		0.123		0.136	

Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1

4.6.1 Movement and transport costs

As mentioned, the LMES 2011 dataset has 277 individuals who moved from their primary sampling unit in the previous year. These observations were removed from the OLS estimations because of the endogeneity concerns discussed at the end of this chapter. However, these observations make up approximately 20% of the sample. These individuals may be relocating to move to an area where transport costs are lower, and thus the probability of finding a job, or net take-home pay after transport costs, is higher. Explicitly examining this allows us to better understand one of the factors which may result in transport costs being endogenous.

In 2011, respondents were asked whether they had moved from their previous location or cluster. To measure the possible effect that this transition has on the data, these individuals need to be analysed in the year before they moved, that is, 2010. A dependent variable (*moved*) was created that captured this movement as a dummy variable and was then included in the 2010 data. All other variables for the 2011 analysis were constructed in the same way for the 2010 data and analysis. The 2010 dataset includes 1860 observations. Table 4.10a indicates that 493 observations are missing (from the total of 1860). These individuals did not respond to the question of whether they had moved in 2011—essentially they are not part of the 2010 and 2011 panels. The analysis concentrates on the employed and unemployed groups, for which two variables were previously constructed (see Table 4.3). Table 4.10a summarises each reported employment status against *moved*. Question 2 (the QLFS version) was less strict, but more accurate, in its definition of unemployment, and therefore includes more observations. Once again the QLFS version is used in this analysis. Table 4.10b summarises the differences in transport costs between people who did and did not move. Apart from the number of observations, no meaningful differences can be seen.

Table 4-10a: Employed or unemployed by moved

	Question 1			Question 2		
Moved	Unemployed	Employed	**Total**	Unemployed	Employed	**Total**
No	657	216	**873**	783	307	**1,090**
Yes	156	61	**217**	195	82	**277**
Total	813	277	**1,090**	978	389	**1,367**

Table 4-10b: Summary statistics of median transports costs by moved

Moved	Mean	Median	Stv.Dev	Min	Max	Obs.
No	462.1525	400	271.4568	40	1300	59
Yes	443.9333	400	200.1922	88	900	15
Total	458.4595	400	257.4638	40	1300	74

Table 4.11 shows a summary of all the variables. Both full-time and part-time work experience were not captured in 2010, and were therefore not included in this analysis.

Table 4-11: Summary statistics of moved *for the unemployed and employed (2010)*

	Moved					
LABELS	Mean	Median	Stv.Dev	Min	Max	Obs.
Median transport cost	409.6	400	139.9	88	800	1,563
Gender: male = 1	0.407	0	0.491	0	1	1,860
Age (constructed)	23.52	23.45	1.560	18.18	31.00	1,855
Matriculation with exemption	0.446	0	0.497	0	1	1,860
Matriculation without exemption	0.361	0	0.480	0	1	1,860
Certificate/Diploma with less than grd12	0.0565	0	0.231	0	1	1,860
Certificate/Diploma with grd12	0.260	0	0.439	0	1	1,860
Bachelor's degree or higher	0.0151	0	0.122	0	1	1,860
Resides in a rural area: rural = 1	0.0893	0	0.285	0	1	1,321
Resides in an informal settlement: informal = 1	0.468	0	0.499	0	1	1,321
Kwazulu-Natal	0.182	0	0.386	0	1	1,860
Limpopo	0.270	0	0.444	0	1	1,860
Constant						
	1,110					

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Base category: female (male=0); no matric; school only; not in rural area; not in informal settlement; sampled from Gauteng

A preliminary estimation of the relationship between those who *moved* and *median_trans* is shown in Figure 4.9. It suggests that there is a negative relationship between moving and transport costs. This means that the higher the transport cost, the less likely individuals from a particular cluster are to move. This could mean that that it is unlikely that these people moved because they were seeking lower transport costs, but rather for other reasons.

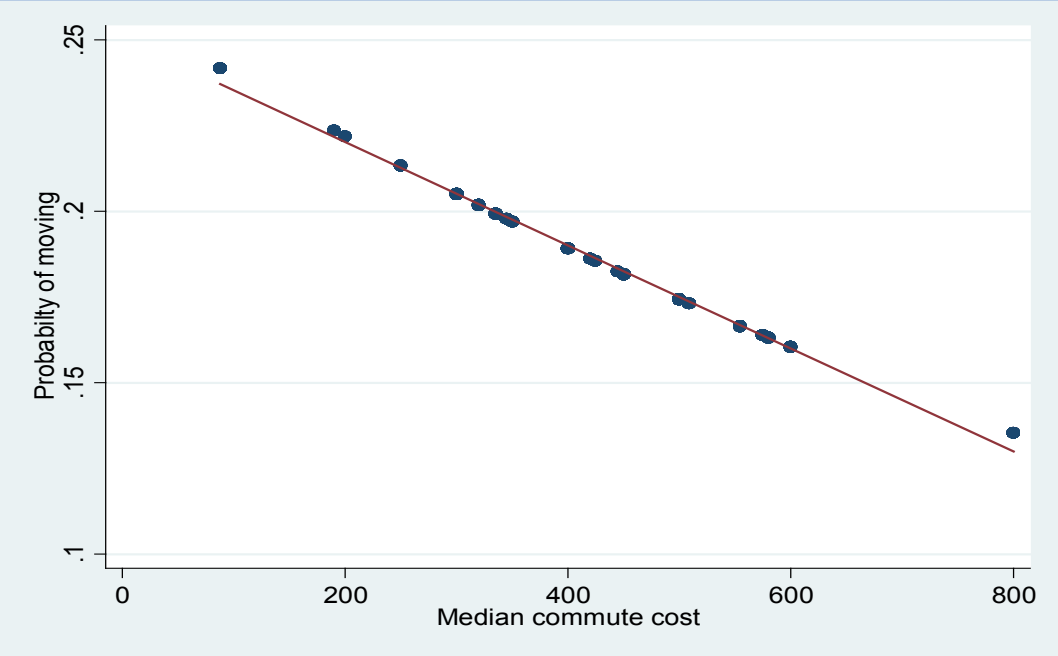


Figure 4-9: Observed and fitted logits by mean commute costs

The final estimation with all covariates in Table 4.12a shows that there is a negative, but insignificant, relationship between transport costs and moving. It also suggests that age and gender do not affect the decision to move. There is also a strong negative relationship between moving and residing in a rural area. This suggests that residing in a rural area decreases the probability of moving, which could mean that transport costs in these areas are such an impediment that they greatly deter moving. Table 4.12b shows the declining probabilities of moving as transport costs rise, starting with a mean value of approximately R410.

Table 4-12a: Logit estimation of moved *for the unemployed and employed*

Variables	Coefficients Moved	Se
Median transport cost	-0.00108	(0.000711)
Gender: male = 1	-0.0160	(0.161)
Age (constructed)	0.0525	(0.0514)
Matriculation with exemption	-0.397**	(0.200)
Matriculation without exemption	-0.871***	(0.232)
Certificate/Diploma with less than grd12	0.444	(0.333)
Certificate/Diploma with grd12	-0.0469	(0.186)
Bachelor's degree or higher	0.914*	(0.511)
Resides in a rural area: rural = 1	-0.614*	(0.370)
Resides in an informal settlement: informal = 1	0.194	(0.181)
Kwazulu-Natal	0.631**	(0.245)
Limpopo	0.322	(0.199)
Constant	-2.051	(1.248)
Observations	1,110	
Chi-squared	0.000105	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Table 4.12b: Move probabilities for median daily commute costs

	Mean	R440	R480	R520	R540
Pr(y=1x):	0.1810	0.1745	0.1683	0.1624	0.1594

Given these results, it is unlikely that any just further conclusions can be made regarding the effect that moving had on the 2011 sample. The possibility that individuals self-select themselves into areas with lower transport costs still exists, but it is unlikely based on these results. Therefore, the 277 observations were excluded from final estimates. However, because of the results, a reservation wage OLS is provided by Table 4.13. When compared to Table 4.8, there is very little change in the significance of the covariates, except for minute changes in the magnitudes. Having a certificate or diploma with less than a Grade 12 qualification gained significance at the 10% level.

Table 4-13: OLS estimation for the employed and unemployed with moved

VARIABLES	lnrwfar coef	se	lnrwnear coef	se
Median transport costs	0.000225**	(0.000110)	0.000135	(0.000117)
Gender: male = 1	0.144***	(0.0325)	0.189***	(0.0344)
Age (constructed)	0.0177*	(0.0105)	0.0138	(0.0111)
Matriculation with exemption	0.383***	(0.0549)	0.408***	(0.0582)
Matriculation without exemption	0.297***	(0.0392)	0.283***	(0.0416)
Certificate/Diploma with less than grd12	0.126*	(0.0678)	0.0728	(0.0721)
Certificate/Diploma with grd12	0.0185	(0.0384)	0.0135	(0.0407)
Bachelor's degree or higher	0.606***	(0.154)	0.648***	(0.159)
Between 3 months and 1 year full-time work	0.00933	(0.0394)	0.0416	(0.0418)
Between 1 and 3 years full-time work	0.126***	(0.0416)	0.202***	(0.0439)
Between 3 months and 1 year part-time work	0.0985***	(0.0362)	0.107***	(0.0384)
Between 1 and 3 years part-time work	0.128***	(0.0485)	0.187***	(0.0513)
Rural_urban	-0.112	(0.0692)	-0.154**	(0.0735)
Informal settlement	-0.0168	(0.0349)	0.00807	(0.0370)
Kwazulu-Natal	-0.0383	(0.0520)	-0.0741	(0.0551)
Limpopo	-0.108**	(0.0418)	-0.133***	(0.0444)
Constant	7.393***	(0.252)	7.181***	(0.267)
Observations	1,165		1,173	
R-squared	0.151		0.170	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Base category: female (male=0); no matric; school only; zero to 3 months full/part time work experience; not in rural area; not in informal settlement; sampled from Gauteng.

Table 4.13 suggests that individuals' reservation wages increase by about 0.0225% for every rand spent per month on commuting to jobs that are far. This amounts to a 2.25% increase in reservation wages for every R100 spent. To compare these results to those in Table 4.8, the constant needs to be considered. The anti-log of the constant amounts to approximately R1635 for all base categories in the regression. This suggests that for every rand spent, individuals expect 36.5 cents in return. This is approximately a 1.5 cent increase from the estimates in Table 4.8. This suggests that when movers are included, the results are biased to higher transport areas. Similar results were found for closer jobs, but with smaller magnitudes.

4.7 Robustness

To verify some of the results, a regression using *desperate wage* against the same independent variables was estimated. Table 4.14 displays the estimated results when using *desperate wages*. It shows that transport costs are not significant. This question was asked as a control to the previous reservation wage questions. The table shows that the desperate wage seems to reflect what would be

an expected reservation wage for the sample areas in terms of actual values. This may be because respondents have removed critical elements from their reservation wage decision. One of these elements could be transport costs. Because they were asked what they would accept in desperation, these respondents may not have included commute costs in their responses. Further evidence of this argument can be found in the coefficients, which are much smaller than their near and far counterparts.

Table 4-14: OLS estimation for the employed and unemployed

VARIABLES	Entire sample desp_wage	se	Entire sample lndesp_wage	se
Median transport cost	-0.298	(0.653)	-3.37e-05	(0.000108)
Gender: male = 1	367.7*	(191.9)	0.124***	(0.0318)
Age (constructed)	81.07	(61.86)	0.0327***	(0.0103)
Matriculation with exemption	731.8**	(322.0)	0.273***	(0.0534)
Matriculation without exemption	612.9***	(232.6)	0.207***	(0.0386)
Certificate/Diploma with less than grd12	33.57	(402.2)	0.0727	(0.0667)
Certificate/Diploma with grd12	-140.4	(226.8)	0.0407	(0.0376)
Bachelor's degree or higher	1,793**	(818.6)	0.458***	(0.136)
Between 3 months and 1 year full-time work	49.73	(234.3)	0.00564	(0.0389)
Between 1 and 3 years full-time work	749.9***	(244.4)	0.144***	(0.0405)
Between 3 months and 1 year part-time work	340.7	(212.8)	0.0568	(0.0353)
Between 1 and 3 years part-time work	313.5	(284.3)	0.119**	(0.0472)
Rural_urban	-522.5	(408.7)	-0.122*	(0.0678)
Informal settlement	270.1	(202.1)	0.00163	(0.0335)
Kwazulu-Natal	-164.8	(378.7)	-0.0869	(0.0628)
Limpopo	-543.9**	(265.9)	-0.208***	(0.0441)
Constant	-534.2	(1,494)	6.470***	(0.248)
Observations	1,141		1,141	
R-squared	0.042		0.134	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.8 Endogeneity and selection bias

A key issue in this analysis which this study cannot fully overcome is that of endogeneity. There are a number of ways that this may occur. There are unobservable, or unaccounted for, characteristics of individuals that are correlated with transport costs and are thus present in the error term. Individuals may also have selected themselves into certain areas because transport costs are lower there, which may result in reverse causality between the outcome variable of interest and transport costs (simultaneity). Given the nature of the data, controlling for endogeneity is difficult. However, I have

tried to deal with some of the mechanisms that may result in endogeneity of transport costs as well as understand selection into the sample and into different clusters in two ways.

The first way is to create cluster wide transport cost variables. This eliminates potential direct correlation between unobservable individual effects and transport costs. The second approach is to directly model the decision to move and its relationship with transport costs, as discussed above. The results obtained have two main implications. Firstly, that movement and transport costs are negatively related. This suggests that individuals in clusters with lower transport costs are more likely to move. Although this does not eliminate the possibility of selection into low transport areas for purposes related to labour market outcomes, it does suggest that people are not moving from high transport cost areas to lower cost areas. Secondly, the results indicate that the sample, once movers are eliminated, is biased towards lower transport cost areas.

4.9 Summary

This chapter discusses the relationship between reservation wages and transport costs by using a Mincerian earnings equation with various other covariates. This was done using median transport costs (essentially monthly costs) on reservation wages for jobs that are far and nearby. A positive and significant relationship was found between median transport costs and reservation wages for jobs that are far. However, these results are primarily driven by the employed sample of the LMES. This may indicate that the unemployed underestimate the impact that transport costs have on their commutes. Although only descriptive, the results in this chapter suggest that individuals try to recoup a portion of their transport costs and those individuals who live closer to job centres recoup less. Transport costs and the decision to move were also modelled to better understand the possible selection bias the sample could have. Although selection bias could not be eliminated, the results from this section suggest that movers are unlikely to relocate on the basis of high transport costs.

5 Chapter 5

5.1 Introduction

This section analyses the relationship between transport costs and the transport premium. This refers to the premium added to wages required for travelling to job search locations that are farther away. Time as a proxy of transport costs is also examined. Time has often been used as a proxy for transport costs, as seen in Arnott (1998), Zenou and Smith (1995) and Zenou and Brueckner (2003). However, this analysis is different, because it uses transport cost premiums in relation to time and transport costs. This is important because proximity to employment locations can affect transport costs and their proxies. Holzer and Ihlanfeldt (1996) use an LPM to assess whether the proximity of firms to both residential locations and public transit points affect black employment and wages in Atlanta, Boston, Detroit and Los Angeles. They use firm level data to do this and find that the probability of employment increases the closer the employer is to the black employee's public transport point and residence. Their findings are consistent with the SMH and also point out that, although black employees live closer to employers, on average their cost per mile to access those jobs is higher than for whites who live in outlying suburban areas. Holzer and Ihlanfeldt (1996) further assert that transportation improvement policies must be designed to improve access to higher wage jobs.

5.2 Methodology

5.2.1 Hypothesis

An OLS regression is used to understand the significance that median transport costs and time have on the transport related component. The transport related component measures the difference between near and far reservation wage responses, called the transport cost premium. The LMES data is used for this purpose. This section indirectly focuses on the reservation wage aspect of this decision by analysing the relationship that exists between transport premiums and transport costs. The LMES provides a unique opportunity to investigate transport premiums because it asks two reservation wage questions successively. The hypothesis is that transport costs do not increase premiums.

5.3 Descriptive analysis (econometric specification)

5.3.1 Dependent variable

Transport premiums are measured as the difference between the minimum wage for a place of work that is far from a place of residence and the minimum wage for a place of work that is nearby a place of residence. Both questions are endogenous because of their subjective nature, which potentially makes them poor spatial (in terms of distance or time) indicators of transport costs. However, by determining the difference between the answers to the first and second questions, it is possible to extract the premium that each individual is willing to pay to travel to a job that is far away while removing some endogeneity. This is plausible because the questions were asked in close succession, so that respondents could better distinguish the difference between the two questions. It is assumed that very little travel is required for with jobs that are nearby and that jobs that are far require significant travel. Van der Berg and Gorter (1997) identify their dependent variable in a similar manner. More importantly, it is likely that any reasonable individual would increase their minimum wage for a more distant job, because of the increased costs associated with that distance. For simplicity, it is assumed that the entire difference is taken to equal transport costs, since the first minimum wage (near to home) includes other common independent costs related to searching. These independent costs could include printing curriculum vitae, buying newspapers, making telephones calls and other related costs. Although the costs associated with searching for jobs closer to home are smaller, it is likely that the bulk of the increased cost accommodates the increased commute distance. In fact, because respondents answered the questions immediately, it is unlikely that they would relate the cost to anything other than the increased distance required to travel to employment that is farther away.

Many studies assume that the entire premium is equivalent to the transport cost. For instance, Arnott (1997) uses a proportion ($\beta - 1$) of the reported wage as the transport cost. The weighting here differs to that used by Patacchini and Zenou (2005), who assert that the unemployed spend less time than the employed on job searching and commuting activities. Wasmer and Zenou (2002) assume that the employed travel into employment centres five times more than the averaged unemployed person partakes in search and commuting related activities. Zenou and Brueckner (2003) follow Smith and

Zenou's (1995) approach to specifying the transportation costs of unemployed individuals, as does this paper. They assert that, to make the analysis manageable, the entire portion of the commuting costs of employed workers is given to unemployed workers. This assumes that unemployed workers commute as frequently as employed workers when searching for work or commuting for search and non-search (leisure) activities. Furthermore, this reduces the variable effect that employment status has on location related costs because individuals residing in the same areas pay similar land rents (Zenou and Brueckner, 2003).

Table 5.1 shows a breakdown of the main dependent variable, *premium*, as it was created. The table shows the means differences and not the actual differences.

Table 5-1: Summary statistics of reservation wage indicators

Stats	*rwfar* *A*	*rwnear* *B*	*premium* *(A – B)*
Mean	4241.473	3231.08	1098.137
p50	3500	2500	800
Sd	2818.717	2832.745	1384.081
Min	500	500	0
Max	25000	60000	15000
N	1356	1359	1342

Figure 5.1 shows that each cluster has two or more values for *premium.* It also reflects large differences between individuals' transport costs. These differences are mainly caused by differences within each gender, cluster or province group. The between group variation is about a fifth of the within group variation. Since the assumption is that this difference forms the commute cost, it seems to suggest that individuals are accessing jobs that vary greatly in distance from homes. A group of individuals from the same cluster might travel short distances to work, while another group is willing to travel far greater distances. Figure 5.2 shows the density of the transport premium. The two lines reflect the necessary adjustments that were implemented to remove nuisance observations. For instance, 340 observations had zero differences and were created if a respondent's difference was negative. Some differences were also deemed too large, for instance R15000. The blue line in Figure 5.2 clearly shows this. The orange line on the same diagram shows the adjusted observations. Figures 5.1 and 5.2 show that there is a large concentration of individuals with a premium of around R1000—

248 observations to be exact. Another 231 respondents had a R500 transport premium. These costs seem to be in line with what is expected in terms of transport costs for searching or working.

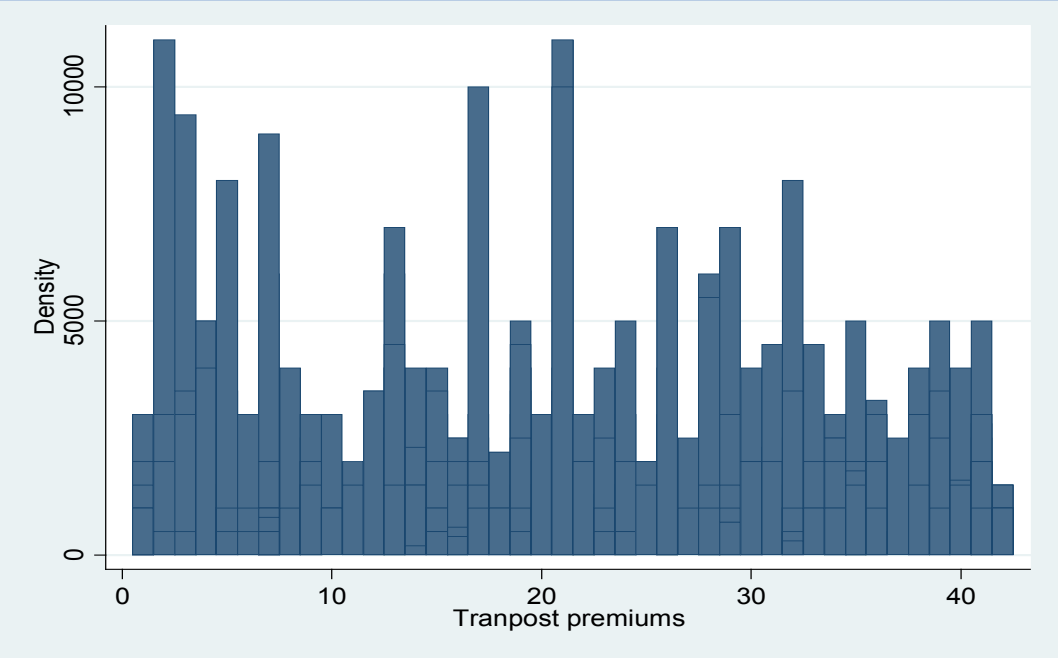


Figure 5-1: Transport premium by cluster

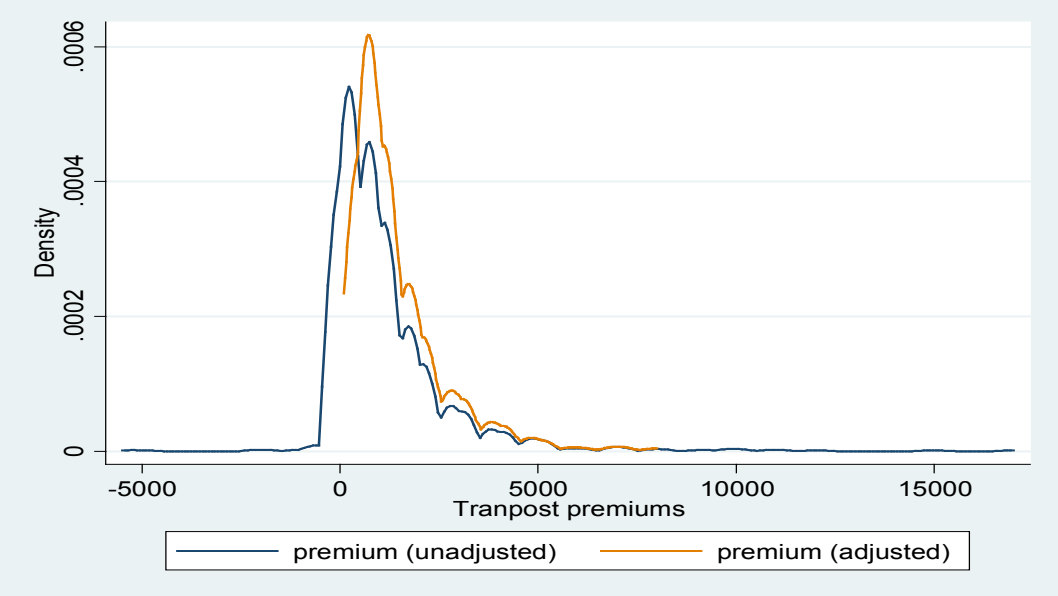


Figure 5-2: Transport premium adjustments

This is a potentially novel idea because I could not indirectly extract transport costs without the individual's knowledge of it. In other words, the potential endogenous effect of transport costs has

been limited. This isolates the commute cost and can then be used to ascertain what relationship it has with the median transport costs reported by the employed portion of the LMES sample. Many outliers were present in these variables, especially on the higher end of the distribution. Some individuals might be willing to bear such high minimum wages if they, for instance, had a reasonably high education. Matching the outliers with their respective education levels shows that some respondents indicated excessive minimum wages for low levels of education. It is possible that some exceptional individuals can acquire wealth despite acquiring little or no education. However, this is highly unlikely given the sample area and population. Therefore, 25 observations were dropped. Furthermore, some observations had higher nearby minimum wages than their far counterparts. This created negative values. Therefore, 14 such observations were removed. The log of *premium* was then taken to create *lnpremium*. Figure 5.3 displays the normal probability plot[9] for *lnpremium* and shows that the variable is quite normally distributed. This could be a result of the way that respondents in the LMES answered the reservation wage questions. Even though the question was asked in a continuous fashion, respondents chose to answer it in ranges, creating what looks like a discrete variable. This caused many of the same minimum wages to be repeated in the sample, indirectly creating a semi-discrete variable. This may be why a zig-zag pattern of the normal values toward the upper end of the distribution is seen in Figure 5.3. The kernel density function of *lnpremium* in Figure 5.4 confirms that the variable is normally distributed because transport premiums cluster around R500 and R1000. The swell of zeros on the left is the result of the log transformation procedure, but has no bearing on the variable itself.

[9] A normal probability plot for any variable is the computed expected normal values created after the variable is ranked and sorted. It plots these values against the actual normal values. The expected normal values follow a normal distribution while the normal values follow the actual distribution. The expected normal plot is the diagonal line and ideally the actual normal values should also be scattered along this line. This means that the data is plotted against a theoretical normal distribution. The vertical axis displays ordered response values and the horizontal axis shows normal order statistic medians.

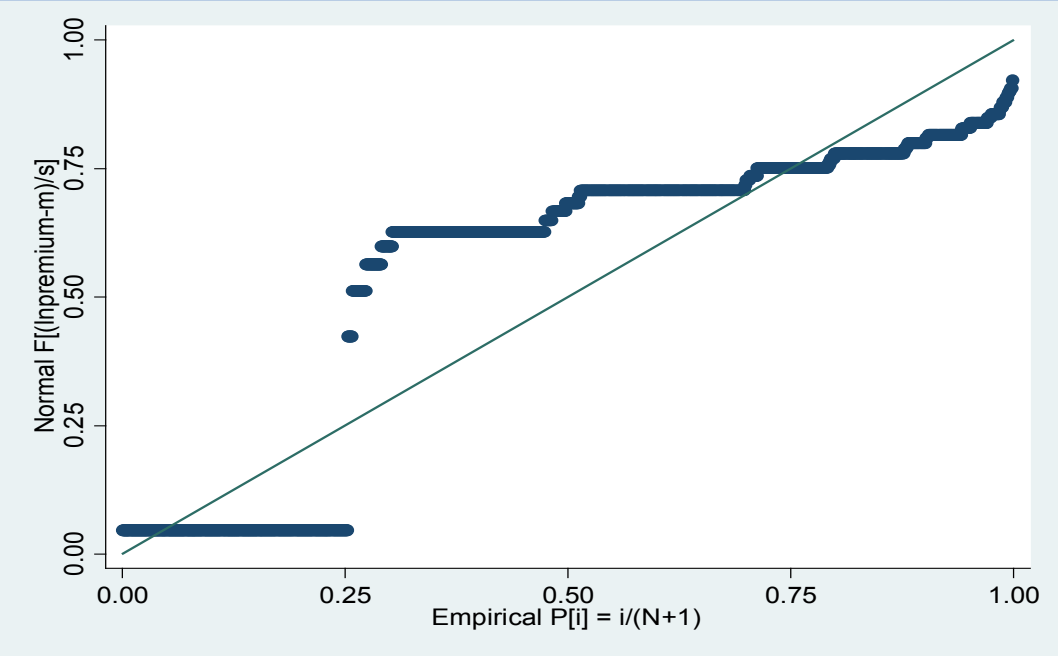


Figure 5-3: Normal probability plot for the log of premiums

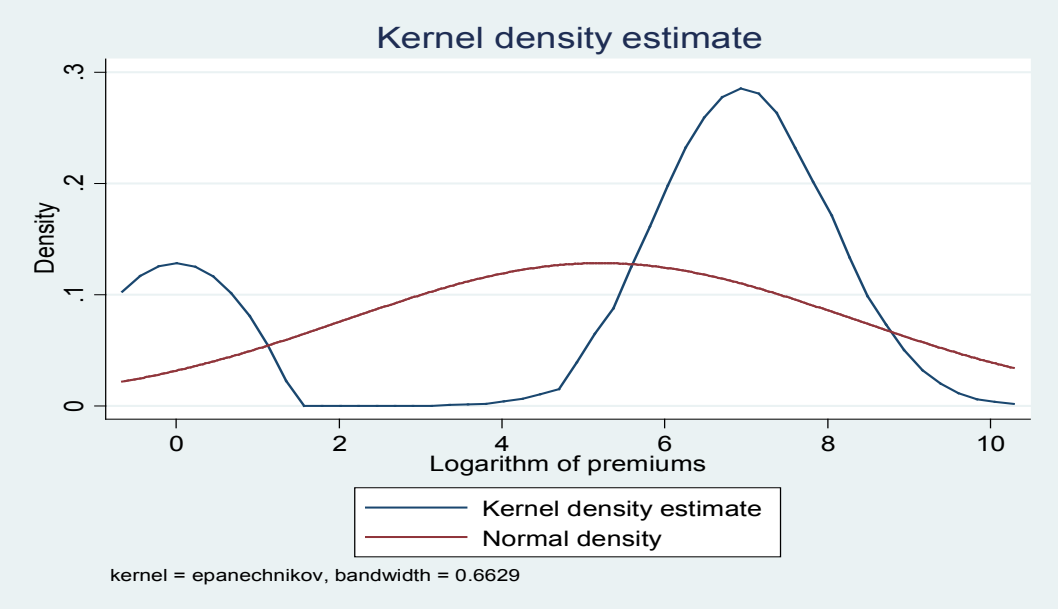


Figure 5-4: Density for the log of premiums

The covariate selection criteria for this analysis were different to the ones used in the previous section. Indicators that affect reservation wages do not necessarily affect transport costs. The size of commute costs is essentially a choice between alternatives. However, it is better to avoid such endogenous variables. Spatial variables were selected instead, as it is less likely that individuals had much choice in terms of the factors that affect their locality.

5.3.2 Independent variables

When considering other transport related or spatial variables, it is ideal to have variables that do not capture information already captured in the cluster variable. Time and distance are good spatial indicators and make good control variables in a regression, because they enhance the analysis by providing another control for the costs involved in commuting. However, the LMES does not have time and distance specific variables. Google Maps was used to get a meaningful measure of distance or time from each respondent's residence cluster to the closest reasonable CBD. The commute distance or time is based on travelling by car and not public transport, which would presumably take longer. This study assumes that the CBD would be the job centre where individuals in each cluster would work, or attempt to find work. Although this is highly presumptuous, because there is no firm level data to offer any indication of the agglomeration in each area, it is a plausible indicator. Google Maps provided 1 to 3 possible routes from each cluster to a CBD. The routes included time and distance estimates. The mean of these values was calculated to create covariates that could be included in the final analysis. Distance and time were highly correlated with each other. Since they are both transport cost proxies, either could be used in the estimation. Arnott (1997) notes that time related costs are more important than the monetary costs of commutes. However, finding the time related cost would require collecting data on all the transport related costs incurred by each mode used by respondents. In South Africa, the most common mode of transport is taxis. Taxis operate in an industry that is, for the most part, not formalised or regulated by government. This means that obtaining standardised estimates for transport costs by taxi route is very difficult given the number of routes which would need to be considered. The mean values for time and distance shown below correspond with what is expected of a common daily commute represented in minutes and kilometres, respectively. Figure 5.5 shows the distribution of mean time and distance for each province. The diagram shows the mean time and distance to a CBD from a central residential location by province. It is particularly interesting that the central values for Kwazulu-Natal are far more spread out than the Limpopo and Gauteng values. This might indicate that the CBDs in this province are far more dispersed, with individuals either living fairly close to CBDs or in remote areas. Gauteng has an even distribution of CBDs that are close and a medium distance away. The graph suggests that individuals

residing in Limpopo are better off, in terms of time and distance of commutes. Figure 5.6 displays the distribution of mean time and distance for the entire sample. It shows that the highest concentrations of values are individuals traveling between 20 and 60 minutes or kilometres per commute to a CBD. Table 5.2 shows the mean values for different intervals of mean time. Importantly, there are mean times of more than an hour, which is where a turning pointing is expected, in terms of the premium and mean time relationship.

Table 5-2: Descriptive statistics for mean time intervals

Mean time	Obs	Mean	Std. dev	Min	Max
0 to 30 min	364	18.24038	6.925761	2	27
31 to 45 min	272	36.96875	5.149601	31	45
46 to 60 min	312	53.38462	3.795134	47	59.5
61 to 75 min	100	65.42	3.601851	61	70
76 to 90 min	135	81.71111	3.75334	76	85
91 min or more	138	110.4348	11.38315	94	133

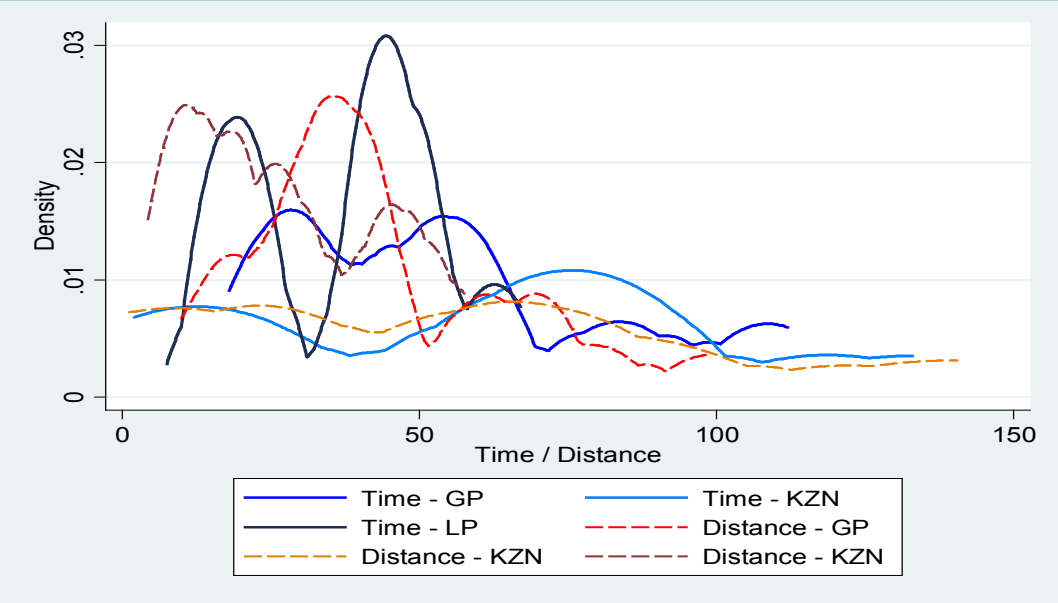


Figure 5-5: Density of mean time and distance by province

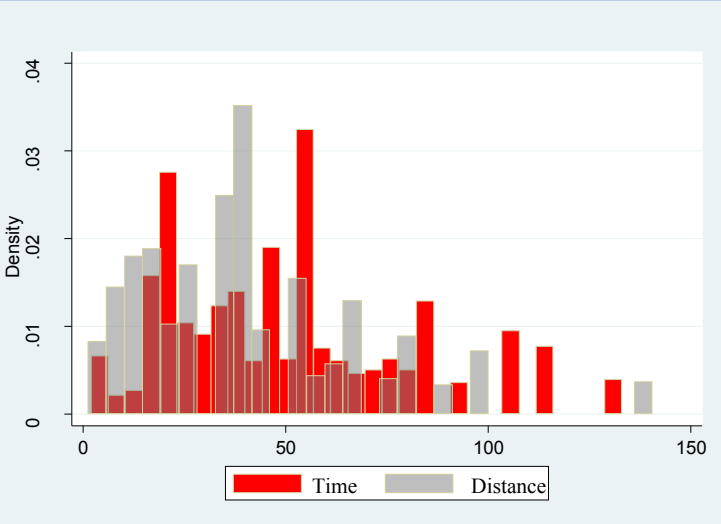


Figure 5-6: Mean time and distance by province

Figure 5.7 displays the symplot for mean time above and below the median. It suggests that there are many outliers. However, it was constructed by cluster, which means that groups of similar values create the staggered effect that is seen. Ultimately, time was chosen instead of distance because the literature usually takes this approach.

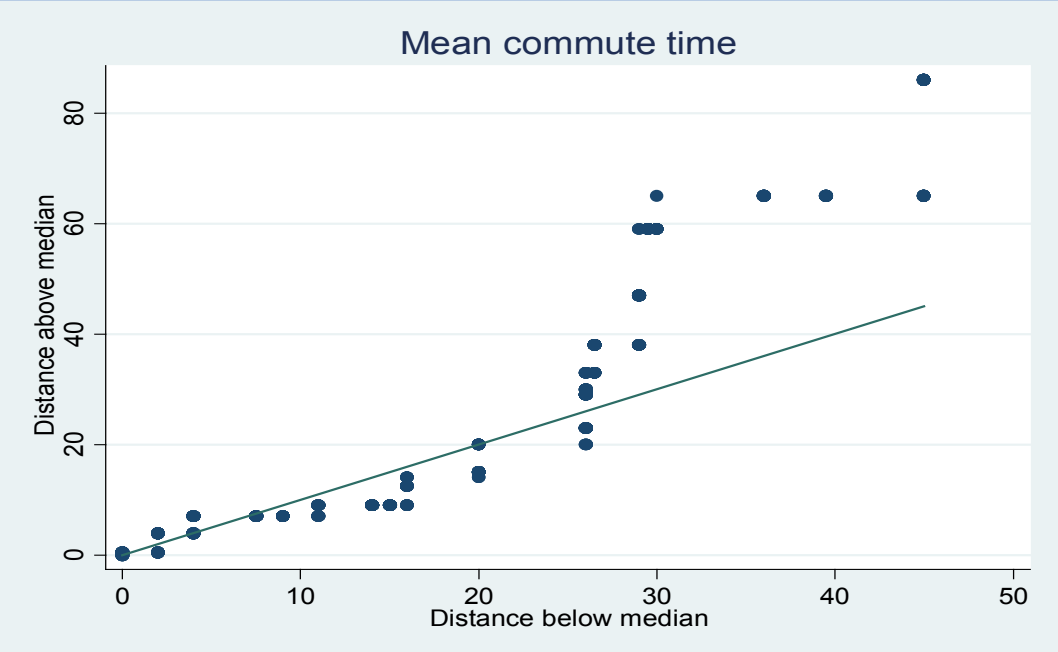


Figure 5-7: Symplot of mean time

Chapter 4 discusses how median transport costs are constructed. The other covariates are the same as those used in Chapter 4 and are used in the final estimations in this section.

5.4 Results

This section aims to understand the relationship between transport premiums and transport costs. An OLS regression was used on *transport premiums* and *transport costs*, together with other covariates, to achieve this aim. Special care needs to be taken when disseminating information about the relationship that exists between median transport costs and transport premiums. This relationship is interpreted differently in this analysis because the dependent variable is a type of transport cost itself—a premium. To do this, the analysis must include some basic urban economic concepts that have been mentioned in this paper. In general, it is expected that individuals who live farther from CBDs, for instance in rural areas, face higher transport costs than their urban counterparts. This is generally compensated by lower land rents. People living in urban areas face the opposite, that is, lower commute costs and higher land rents. This means that median transport costs might be measuring, or acting as a proxy, for lands rents, in that people who are posting high transport premiums have declining land rents. This is because, people who are increasingly farther from CBDs, presumably living in non-urban areas, have steadily declining land rents,. However, they inevitably post higher premiums to compensate for the increased commutes. This example describes a negative relationship between transport costs and premiums. Van der Berg and Gorter's (1997) work is used to understand a positive relationship between transport premiums and median transport costs. They construct their dependent variable in a similar fashion and find that people in urban areas are more willing to pay for commutes than their rural counterparts. This, like the transport cost and transport premium relationship, might suggest that people living in urban areas can incorporate personal and other leisure activities into their trips or that traffic congestion increases their transport costs. At every level of the premium, more of these people are willing to pay for commutes.

Table 5.3 shows the final regression output for this analysis. It includes three separate regressions for the entire sample, the employed sample and the unemployed sample. The results suggest that the significance of median transport costs is primarily driven by the employment sample. This is not surprising because it is also seen in Chapter 4, which shows that most of the significance generated was also reflected by the employed groups. This can be seen from the second column which shows that a negative relationship exists between median transport costs and premiums. Therefore, every

rand increase in median transport costs results in a decrease of approximately 0.337% in transport premiums. If median transport costs are thought of as a proxy for land rents, this suggests that as land rents decline, reported premiums rise. Also of interest are mean time and its square, *mtime* and *sqrmtime*, respectively, since they expose the relationship between time and transport premiums. The results show that both variables are positive and significant when the employed and unemployed groups are isolated. The coefficient of the mean commute time squared is negative. Together with a positive mean commute time, this suggests that a rise in commute times increases transport premiums at a decreasing rate. Therefore, creating a quadratic relationship between time and this premium indicates that individuals are willing to commit time to commuting to places of work, but only up to a certain critical point (the turning point). At this point, they decide that commuting is no longer feasible or beneficial and stop. Every minute taken to travel increases the premium by approximately 3.8%. This occurs at a decreasing rate, and at around 79 minutes, individuals decide that further travel is no longer worth the employment opportunities that it provides. A visual representation of this relationship can be seen in Figure 5.8. Residing in an informal settlement seems to decrease the commute premium, which might indicate that staying in these areas is a barrier to travel in terms of employment.

Table 5-3: OLS estimation of transport premium

VARIABLES	Both lnpremium	se	Employed lnpremium	se	Unemployed lnpremium	se
Median transport costs	8.86e-05	(0.000666)	-0.00337**	(0.00167)	0.000554	(0.000734)
Mean commute time	0.0381***	(0.0140)	0.0658**	(0.0275)	0.0257	(0.0162)
Mean commute time squared	-0.000236**	(0.000106)	-0.000343*	(0.000207)	-0.000176	(0.000123)
Gender: male = 1	0.175	(0.216)	0.568	(0.410)	0.164	(0.262)
Age (constructed)	0.0705	(0.0691)	-0.0451	(0.136)	0.0741	(0.0810)
Matriculation with exemption	0.503	(0.384)	1.026	(0.714)	0.315	(0.463)
Matriculation without exemption	0.431*	(0.258)	0.930*	(0.539)	0.379	(0.295)
Certificate/Diploma with less than grd12	0.826*	(0.454)	2.418**	(1.104)	0.473	(0.497)
Certificate/Diploma with grd12	0.119	(0.256)	0.786*	(0.465)	-0.193	(0.306)
Bachelor's degree or higher	0.701	(1.308)	2.799	(1.924)	-0.516	(1.823)
Between 3 months and 1 year full-time work	0.276	(0.263)	-0.363	(0.514)	0.485	(0.312)
Between 1 and 3 years full-time work	0.110	(0.275)	-1.106**	(0.497)	0.965***	(0.361)
Between 3 months and 1 year part-time work	0.0239	(0.241)	-0.00840	(0.459)	0.0456	(0.291)
Between 1 and 3 years part-time work	-0.130	(0.317)	-0.505	(0.572)	0.152	(0.391)
Rural_urban	0.486	(0.501)	0.312	(0.961)	0.665	(0.590)
Informal settlement	-0.468**	(0.233)	-0.451	(0.480)	-0.483*	(0.269)
Kwazulu-Natal	0.456	(0.420)	-0.149	(0.869)	0.337	(0.484)
Limpopo	0.178	(0.310)	-0.311	(0.641)	0.288	(0.354)
Constant	1.985	(1.718)	4.812	(3.466)	2.160	(2.000)
Observations	892		254		638	
R-squared	0.027		0.132		0.031	

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1
Base category: female (male=0); no matric; school only; zero to 3 months full time work experience; not in rural area; not in informal settlement; sampled from Gauteng

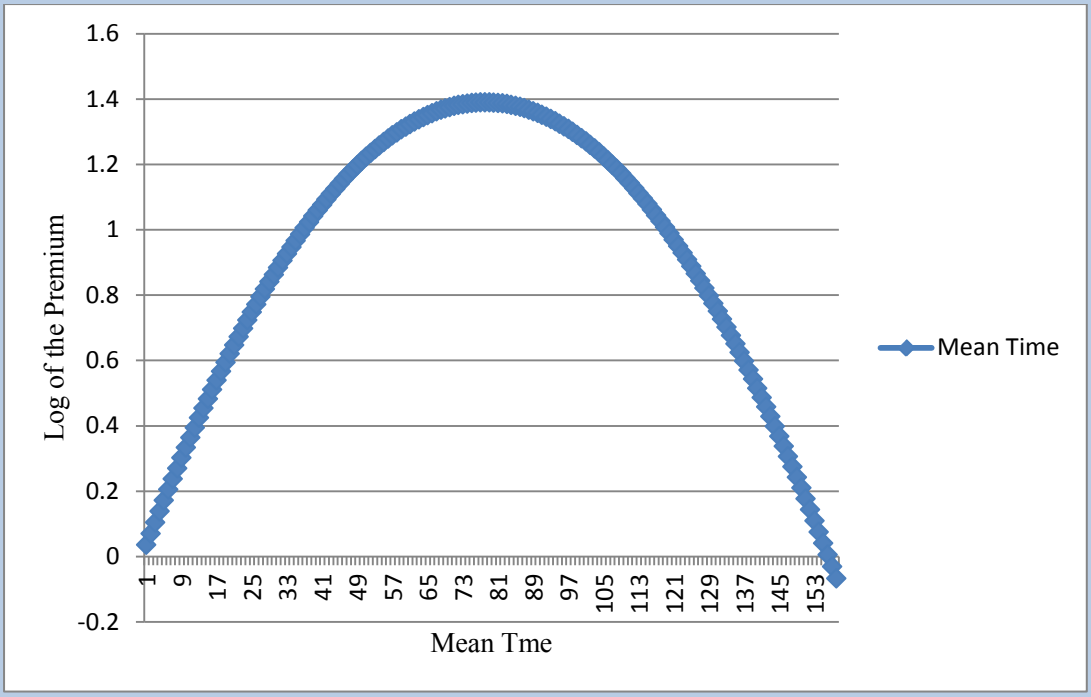


Figure 5-8: Relationship between mean time and transport premium

5.5 Summary

This chapter aims to find a unique relationship between transport premiums and transport costs, and explores time as another proxy for transport costs. This was done using an OLS regression and some evidence is found to show that a negative relationship exists between median transport costs (monthly costs) and transport premiums. The quadratic relationship between mean time and these premiums is also statistically significant. The relationship between transport costs and premiums seems unintuitive. However, when the transport costs are considered as proxies for land rents, premiums rise as land rents decline. This is because individuals living far from work (low land rents) would presumably incur higher transports, which is reflected in their premiums.

6 Conclusions

This paper analyses three different relationships between unemployment and transport costs. First, Chapter 3 examines transport costs and their potential effect on unemployment through searching by building on the work of Kingdon and Knight (2006). It finds evidence that transport costs positively affect search outcomes for unemployed individuals. These results also strongly suggest that the

prevailing unemployment rate in each area significantly decreases the search probability of individuals living in the same or neighbouring areas. Second, Chapter 4 investigates the relationship between transport costs and reservation wage determination using a Mincerian earnings equation. It finds that increased transport costs result in individuals reporting higher reservation wages. Chapter 4 also discusses the effect that individuals moving out of their PSU (cluster) could have on the data and the accompanying results. No meaningful differences in the estimations occurred when these individuals were included in the sample. Third, Chapter 5 explores the unique relationship between transport cost premiums and transport costs, while also focuses on time as a proxy for transport costs. The evidence suggests that transport premiums rise as individuals move farther away from employment centres or urban areas. A quadratic relationship is also found between transport premiums and time (commute time to CBDs), which could mean that individuals stop searching when they reach some critical point in their commute time.

The overarching sentiment of these results suggests that transport costs are a potential barrier to employment outcomes for many South Africans. This evidence suggests that this area of study requires more attention and that increased research in this field could contribute to understanding the relevance that transport costs and transport policy have on the South African economy.

7 Appendix

Table A1: Rural and urban categories for each enumeration area

Gauteng	Urban	Rural	Informal Settlement	Township
Alexandra	•		•	•
Diepkloof Zone 6	•			•
Dlamini	•			•
Eastbank	•			•
Hillbrow	•			
Ivory Park (two clusters)	•		•	•
Jabulani	•			•
Johannesburg CBD	•			
Meadowlands Zone 8	•			•
Mofolo Central	•			
Naledi Ext 1	•			•
Orange Farm Proper		•	•	•
Orlando East	•			•
Protea Glen	•			•
Thulani	•		•	•
Tembisa	•			•
Primrose	•		•	
Lenasia South	•		•	•
Kathlehong	•			•
TokozA	•			•
KwaZulu-Natal				
Durban CBD	•			
Kwadabeka R	•		•	•
Kwambiza	•		•	•
Mozambique B	•		•	•
Qhodela		•	•	
South Beach	•			
Sydenham	•			
Umbumbulu		•	•	•
Umlazi	•		•	•
Waterloo	•			•
Limpopo				
Dikgale (four clusters)		•	•	
Luthuli	•		•	
Madiba Park	•			
Seshego (three clusters)	•		•	•
Makhado	•		•	•
Makwarela	•			•
Lebowakgomo	•		•	•

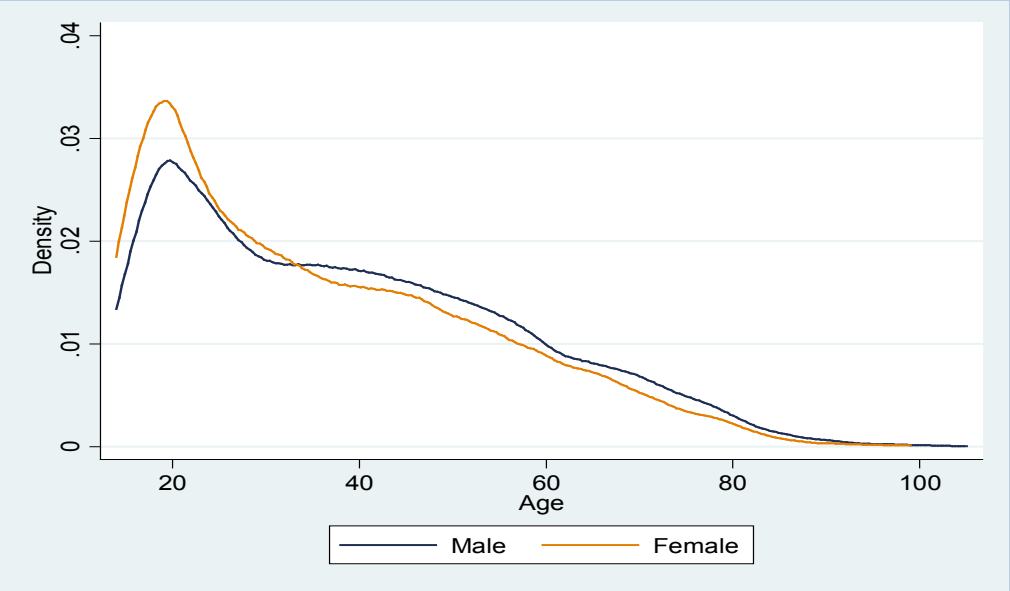


Figure A1: Distribution of age by gender (NIDS)

Table A2: Weekly transport costs by gender

Gender	Mean	Median	Std.Dev	Min	Max	Obs.
Female	123.8866	30	450.1047	0	7000	644
Male	73.22331	20	218.6811	0	5000	1124
Total	91.6776	20	323.5891	0	7000	1768

Figure A2 shows the distribution of education for the sample of unemployed and discouraged people. Only 9% of the sample have a tertiary education , and of that approximately 22% are discouraged.

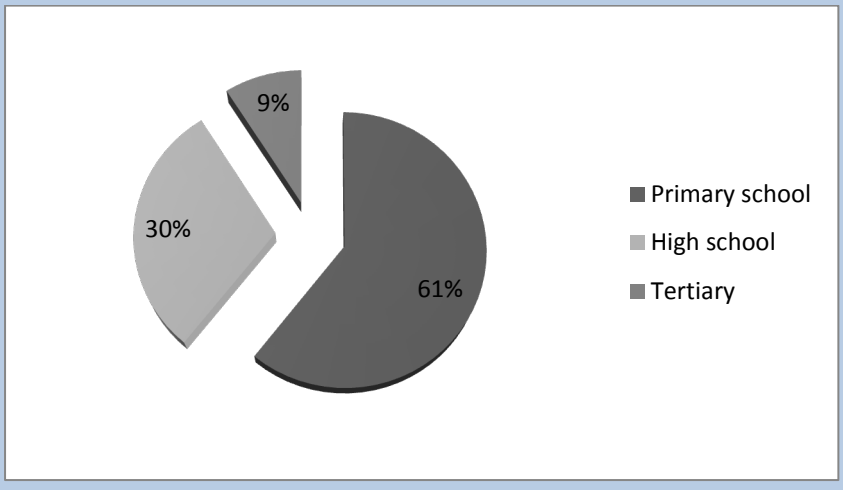


Figure A2: Distribution of education levels (NIDS)

Figures A3 and A4 are density graphs. They clearly show that for poorer households, for both searchers and the discouraged, income is largely based on government transfers. It also shows that as household incomes increase, so does the dependency on government transfers.

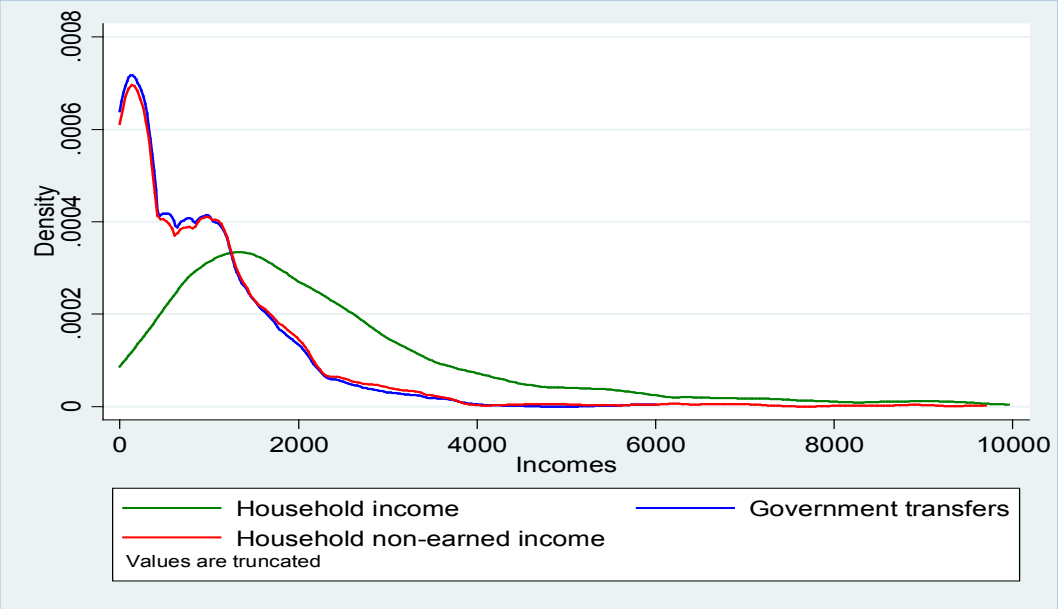


Figure A3: Distribution of incomes for the discouraged

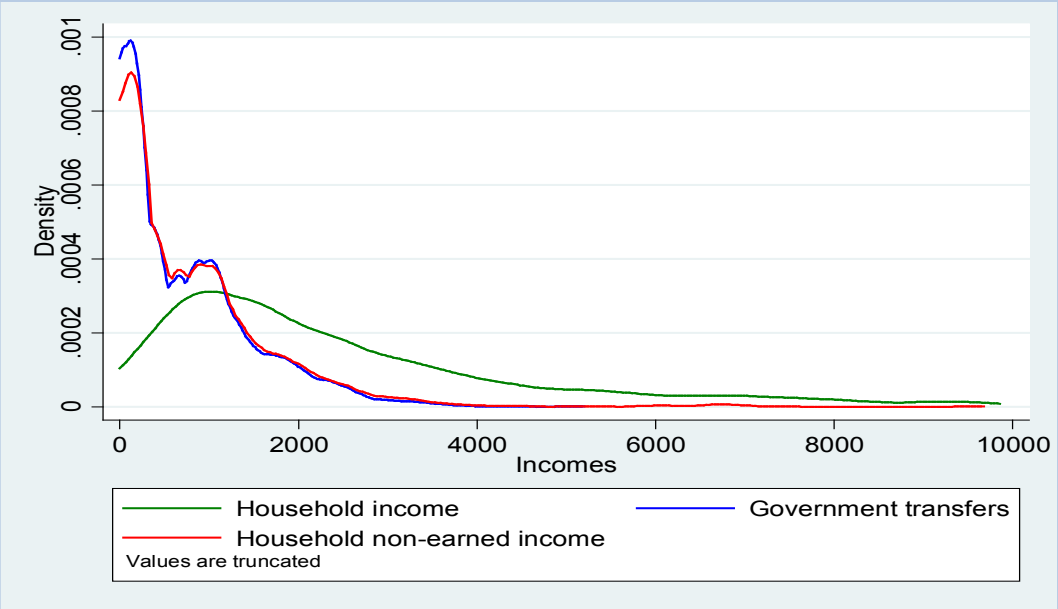


Figure A4: Distribution of incomes for searchers

Figure A5 shows the number of employed members in each household in which an unemployed respondent resides. Most have zero employed members.

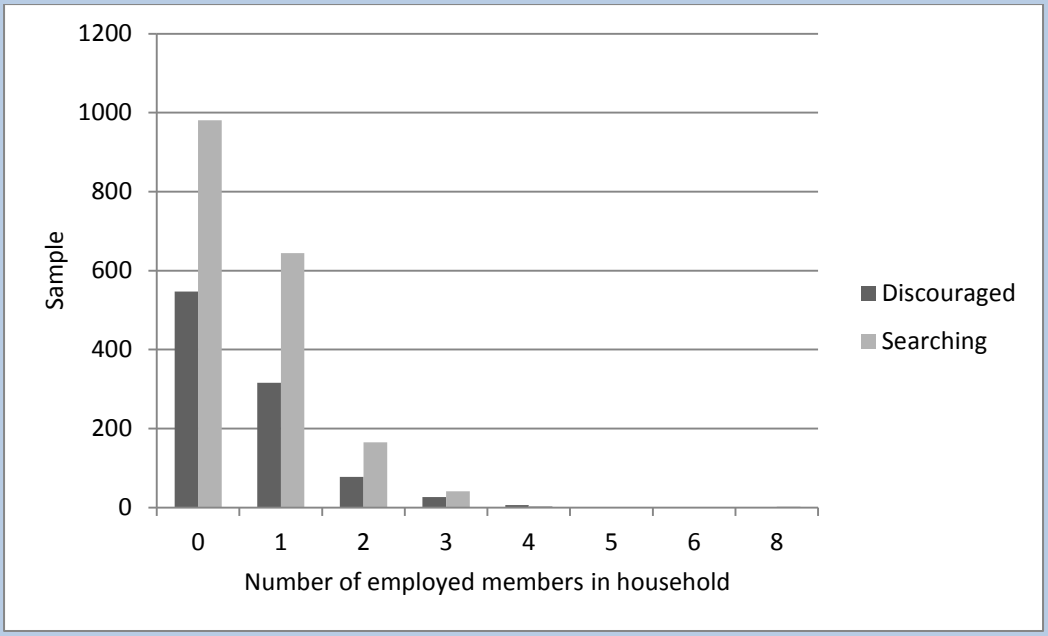


Figure A5: Number of employed members per household

Table A 3: OLS estimation of log reservation wages based on LMES definitions

VARIABLES	Entire sample lnrwfar	Entire sample lnrwnear	Employed & unemployed lnrwfar	Employed & unemployed lnrwnear	Employed lnrwfar	Employed lnrwnear	Unemployed lnrwfar	Unemployed lnrwnear
Median transport costs	0.000228**	0.000141	0.000196*	0.000126	0.000559**	0.000613***	7.37e-05	-4.07e-05
	(0.000112)	(0.000117)	(0.000114)	(0.000122)	(0.000218)	(0.000236)	(0.000131)	(0.000137)
Gender: male = 1	0.146***	0.195***	0.136***	0.185***	0.119*	0.138**	0.0952**	0.154***
	(0.0366)	(0.0384)	(0.0374)	(0.0399)	(0.0606)	(0.0654)	(0.0477)	(0.0497)
Age (constructed)	0.0192	0.0170	0.0175	0.0164	0.00204	0.0110	0.0237	0.0175
	(0.0117)	(0.0123)	(0.0119)	(0.0127)	(0.0200)	(0.0217)	(0.0145)	(0.0151)
Matriculation with exemption	0.439***	0.449***	0.329***	0.343***	0.402***	0.335***	0.177*	0.234**
	(0.0651)	(0.0682)	(0.0704)	(0.0751)	(0.109)	(0.118)	(0.0940)	(0.0986)
Matriculation but no exemption	0.271***	0.249***	0.274***	0.253***	0.350***	0.302***	0.214***	0.203***
	(0.0437)	(0.0460)	(0.0442)	(0.0473)	(0.0785)	(0.0853)	(0.0529)	(0.0555)
Certificate/Diploma with less than grd12	0.103	0.0160	0.174**	0.0874	0.172	0.0510	0.196**	0.126
	(0.0775)	(0.0818)	(0.0790)	(0.0846)	(0.155)	(0.168)	(0.0894)	(0.0938)
Certificate/Diploma with grd12	0.0278	0.0165	0.0838*	0.0476	0.116*	0.0980	0.0697	0.0209
	(0.0436)	(0.0457)	(0.0444)	(0.0474)	(0.0696)	(0.0751)	(0.0563)	(0.0588)
Bachelors degree or higher	0.574**	0.535**	0.568**	0.483**	0.648**	0.531	0.365	0.480
	(0.223)	(0.219)	(0.237)	(0.233)	(0.312)	(0.339)	(0.362)	(0.315)
3 months and 1 year full time work	0.0361	0.0458	0.0498	0.0626	0.0366	0.0613	-0.0109	-0.000231
	(0.0446)	(0.0470)	(0.0447)	(0.0478)	(0.0770)	(0.0835)	(0.0552)	(0.0578)

1 & 3 years full time work	0.137***	0.183***	0.151***	0.197***	0.0537	0.154*	0.0851	0.0523
	(0.0466)	(0.0488)	(0.0465)	(0.0496)	(0.0731)	(0.0790)	(0.0656)	(0.0686)
3 months & 1 year part time work	0.103**	0.0988**	0.0994**	0.0879**	0.0724	0.0961	0.0676	0.0371
	(0.0409)	(0.0430)	(0.0413)	(0.0442)	(0.0688)	(0.0745)	(0.0516)	(0.0541)
1 and 3 years part time work	0.152***	0.191***	0.152***	0.189***	0.0647	0.207**	0.173**	0.139*
	(0.0541)	(0.0566)	(0.0547)	(0.0581)	(0.0856)	(0.0920)	(0.0718)	(0.0750)
Rural_urban	-0.117	-0.222**	-0.0473	-0.166*	-0.0661	-0.142	-0.0883	-0.249**
	(0.0851)	(0.0896)	(0.0854)	(0.0912)	(0.151)	(0.163)	(0.102)	(0.107)
Informal settlement	-0.0302	-0.0162	-0.0454	-0.0341	-0.0411	-0.0136	-0.0276	-0.0171
	(0.0392)	(0.0412)	(0.0401)	(0.0429)	(0.0689)	(0.0746)	(0.0485)	(0.0508)
Kwazulu-Natal	-0.00306	-0.00112	-0.0675	-0.0556	-0.0611	-0.0337	-0.0601	-0.0528
	(0.0699)	(0.0732)	(0.0717)	(0.0762)	(0.136)	(0.143)	(0.0825)	(0.0866)
Limpopo	-0.120**	-0.147***	-0.126**	-0.137**	-0.0748	-0.0524	-0.140**	-0.169***
	(0.0487)	(0.0512)	(0.0513)	(0.0548)	(0.0892)	(0.0969)	(0.0612)	(0.0639)
Constant	7.346***	7.131***	7.383***	7.141***	7.708***	7.169***	7.302***	7.204***
	(0.281)	(0.295)	(0.285)	(0.304)	(0.485)	(0.525)	(0.347)	(0.362)
Observations	902	909	805	811	303	306	502	505
R-squared	0.157	0.170	0.167	0.167	0.186	0.169	0.118	0.115

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Table A 4: Logit estimation of premium based on LMES definitions

VARIABLES	Both lnpremium	Employed Lnpremium	Unemployed lnpremium
Median transport costs	-6.93e-05	-0.00139	0.000211
	(0.000693)	(0.00135)	(0.000818)
Mean commute time	0.0467***	0.0875***	0.0312*
	(0.0146)	(0.0253)	(0.0182)
Mean commute time squared	-0.000300***	-0.000527***	-0.000222
	(0.000110)	(0.000187)	(0.000137)
Gender: male = 1	0.218	0.421	0.178
	(0.225)	(0.363)	(0.296)
Age (constructed)	0.0128	0.0495	-0.0162
	(0.0715)	(0.121)	(0.0896)
Matriculation with exemption	0.525	1.051	0.157
	(0.425)	(0.662)	(0.579)
Matriculation but no exemption	0.476*	0.642	0.421
	(0.266)	(0.471)	(0.327)
Certificate/Diploma with less than grd12	0.922*	1.949**	0.464
	(0.472)	(0.922)	(0.550)
Certificate/Diploma with grd12	0.319	0.405	0.247
	(0.266)	(0.416)	(0.347)
Bachelors degree or higher	2.042	1.961	2.293
	(1.416)	(1.866)	(2.221)
Between 3 months and 1 year full time work	0.192	-0.216	0.380
	(0.269)	(0.462)	(0.340)
Between 1 and 3 years full time work	-0.0140	-0.953**	0.898**
	(0.280)	(0.439)	(0.408)
Between 3 months and 1 year part time work	0.101	-0.0436	0.117
	(0.249)	(0.416)	(0.320)
Between 1 and 3 years part time work	0.0185	-0.658	0.367
	(0.328)	(0.510)	(0.442)
Rural_urban	0.656	0.147	0.940
	(0.513)	(0.905)	(0.634)
Informal settlement	-0.386	-0.591	-0.315
	(0.243)	(0.419)	(0.303)
Kwazulu-Natal	0.371	0.706	0.205
	(0.441)	(0.828)	(0.525)
Limpopo	0.174	-0.125	0.334
	(0.331)	(0.573)	(0.409)
Constant	3.059*	1.726	4.066*
	(1.780)	(3.069)	(2.215)
Observations	796	299	497
R-squared	0.036	0.115	0.034

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

8 References

Albrecht, J. W. and Axell, B. (1984). "An Equilibrium Model of Search Unemployment." Journal of Political Economy, 92(51), 824–840.

Arora V. and Ricci L.A. (2005). "Unemployment and the labour market." In: Post-apartheid South Africa: the first ten years, IMF

Arnott, R. (1998). "Economic Theory and the Spatial Mismatch Hypothesis." Urban Studies Journal Limited

Bantustan territories. In Encyclopedia Britannica Inc. Retrieved from: http://www.britannica.com/bps/media-view/129344/1/0/0

Bhatti, S. H. (2012). "Estimation of the Mincerian Wage Model Addressing its Specification and Different Econometric Issues". (Doctoral dissertation, Université de Bourgogne).

Bhorat, H. (2004). "Labour market challenges in post-apartheid South Africa." South African Journal of Economics, Vol. 72:5

Blanchflower, D. G., and Oswald, A. J. (1995). "An introduction to the wage curve." The Journal of Economic Perspectives, 9(3), 153-167.

Brueckner, J.K., and Zenou, Y. (2003). "Space and unemployment: the labor-market effects of spatial mismatch." Journal of Labor Economics 21, 242–266.

Burdett, K. (1978). "A Theory of Employee Job Search and Quit Rates," American Economic Review, 68(1), 212–220.

Burdett, K., and Mortensen, D. T. (1998). "Wage Differentials, Employer Size and Unemployment." International Economic Review, 39(May), 257–273.

Burda, M., and Profit, S. (1996). "Matching Across Space: Evidence on Mobility in the Czech Republic." Labour Economics, 3, 255–278.

Carolina, M., & Pau, L. (2008). “The Reservation Wage of Unemployed Job Seekers in Curaçao.” Netherlands Central Bank, BNA,WP/08/3

Cichello, P. L., Fields, G. S., Leibbrandt, M. (2005) “Earnings and Employment Dynamics for Africans in Post-apartheid South Africa.” Journal of African Economies, Volume 14, Number 2, pp. 143-190

Clark, W.A.V., Y. Huang, and S. Whithers. 2003. Does commuting distance matter? Commuting tolerance and residential change.Regional Science and Urban Economics 33: 199-221.

Cleveland, W. S. and R. McGill. (1984). “The many faces of a scatterplot.” Journal of American Statistical Association 79: 807 – 822.

Church, A., Frost, M., & Sullivan, K. (2000). “Transport and social exclusion in London.” Transport Policy, 7(3), 195-205.

Development Policy Research Unit. (2012). “Monitoring The Perfomance of The South Afican Labour Market – An Overview of the Youth Labour Market Since 2008.” DPRU-EPP Factsheet 7, August 2012

Diamond, P. A. (1981). “Mobility costs, frictional unemployment, and efficiency.” The Journal of Political Economy, 798-812.

Forster, C. (1983). “Spatial organisation and local unemployment rates in Metropolitan Adelaide: significant issue or spatial fetish?” Australian Geographical Studies, 21(1), 33-48.

Frederick, S., Loewenstein, G., & O'donoghue, T. (2002). “Time discounting and time preference: A critical review.” Journal of economic literature, 40(2), 351-401.

Gibillon, L. Selod, H. and Zenou, Y. (2007). “The Mechanisms of Spatial Mismatch.” Laboratoire d’Economie Appliquee (LEA), LEA-WP 0701

Hall, R. E. (1979). “A Theory of Natural Unemployment Rate and the Duration of Employment.” Journal of Monetary Economics, 5, 153–169.

Hinks, T. (2008). “Poverty Networks and Location Determinants of Job Search in SA.” Journal of International Development, J. Int. Dev. 20, 117–131

Holzer, H. J. (1991). “The spatial mismatch hypothesis: What has the evidence shown?” Urban Studies, 28(1), 105-122.

Holzer, H. J., and Ihlanfeldt, K. R. (1996). “Spatial factors and the employment of blacks at the firm level.” Institute for Research on Poverty, University of Wisconsin-Madison.

Hou, L., Wang, X. and Yu, X. (2011). “Rational Expectation and Education Rewarding: The Case of Chinese Off-Farm Wage Employment.” 2011 International Congress, August 30-September 2, 2011, Zurich, Switzerland, European Association of Agricultural Economists

Jun, M. J., and Hur, J. W. (2001). “Commuting costs of “leap-frog” newtown development in Seoul.” Cities, 18(3), 151-158.

Klasen, S. and Woolard, I. (2000). “Surviving Unemployment without State Support: Unemployment and Household Formation in South Africa.” IZA Discussion Paper No. 237, Bonn.

Kingdon, G. and Knight, J. (2006). “The measurement of unemployment when unemployment is high.” Labour Economics 13: 291 – 315.

Killingsworth, M. R. (1983). “Labor Supply.” Cambridge: Cambridge University Press.

Laird, J. (2006). “Commuting costs and their impact on wage rates.” Institute of Transport Studies, University of Leeds, Working Paper 587

Leibbrandt, M., Mlatsheni, C., van der Berg, S., Burger, R. (2002). “Migration and the changing rural-urban interface in South Africa: What can we learn from census and survey data?” unpublished

Lippman, S. A., and McCall, J. J. (1976). “The Economics of Job Search: a Survey.” Economic Inquiry, 1976a, 14, 347-68.

Louw, A. (2010). “Group Areas: Micro-Histories in the Western Cape. A Bibliography.” African Studies Library, University of Cape Town Libraries, March 2010

Lucas Jr, R., & Prescott, E. C. (1974). “Equilibrium search and unemployment.” Journal of Economic Theory, 7, 188-209.

Mabin, A. (1992). Comprehensive segregation: the origins of the Group Areas Act and its planning apparatuses. Journal of Southern African Studies, 18(2), 405-429.

McCall, John J. (1965). “The Economics of Information and Optimal Stopping Rules.” Journal of Business, July 1965, 300-3 17.

Mincer, J. A. (1974). Schooling and earnings. In: “Schooling, experience, and earnings”. (pp. 41-63). Columbia University Press.

Mortensen, D. T. (1986). “Job Search and Labour Market Analysis.” In: Handbook of Labor Economics, ed. by O. Ashenfelter, and R. Layard, vol. 2, chap. 15, pp. 849–919. Elsevier Science B.V., Amsterdam.

Nattrass, N. (2000a). “The debate about Unemployment in the 1990s.” Studies in Economics and Econometrics, Vol.24, no.3

Nattrass, N. (2000b). “Wage-Strategies and Minimum Wages in Decentralised Regions: The Case of the Clothing Industry in Phuthaditjhaba, South Africa.” International Journal of Urban and Regional Research 24(4) (December): 873-888.

Ong, P. M., and Miller, D. (2005). “Spatial and transportation mismatch in Los Angeles.” Journal of Planning Education and Research, 25 (1),43–56

Patacchini, E., & Zenou, Y. (2005). Spatial mismatch, transport mode and search decisions in England. Journal of Urban Economics, 58(1), 62-90.

Pissarides, C. A. (1979). “Job matchings with state employment agencies and random search.” The Economic Journal, 818-833.

Richardson, H. W. (1988). “Monocentric vs. Polycentric Models.” Annals of Regional Science 22, 2: 1-12.

Robins, P. A. (1981). “The regional impact ofstructural change in Australian manufacturing industry: the case of South Australia.” In: Linge, G. J. R. and McKay, J. (eds) Structural Change in Australia: Some Spatial and Organisational Responses, Canberra, ANU Press, 189-2 18.

Rogerson, R., Shimer, R., & Wright, R. (2005). “Search-theoretic models of the labor market: A survey.” Journal of Economic Literature, 43(4), 959-988.

Rouwendal, J. (2004). “Search Theory and Commuting Behavior.” Wiley online Library, Growth and Change, Vol.35, no.3, pp. 391-418

Rouwendal,J ., and Rietveld, P. (1994a). "A Structural Model of Commuting Distances and Spatial Job Search." Working paper, Wageningen University Dept. of Household and Consumer Studies.

Rouwendal,J ., and Rietveld,P . (1994b). "Changes in Commuting Distances of Dutch Households." Urban Studies, 31, 1545-1557.

Schoer, V. and Leibbrandt, M. (2006) “Determinants of job search strategies: Evidence from the Khayelitsha/Mitchell's Plain survey.” South African Journal of Economics 74(4): 702-724.

Seekings, R., Leibbrandt, M., Nattrass, N. (2004), ‘Income Inequality After Apartheid’, Centre of Social Science Research (UCT), CSSR Working Paper No. 75

Singell, L. D., & Lillydahl, J. H. (1986). “An empirical analysis of the commute to work patterns of males and females in two-earner households.” Urban Studies,23(2), 119-129.

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

Small, K.A., Winston, C., and Yan, J. (2005). "Uncovering the distribution of motorists' preferences for travel time and reliability." Econometrica 73 (4), 1367–1382.

Stiglitz, J. (2000). "The Contributions of the Economics of Information to Twentieth Century Economics." The Quarterly Journal of Economics, November 2000.

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

White, M. J. (1988) "Urban Commuting Journeys Are Not 'Wasteful'." J.P.E. 96 (October 1988): 1097-1110. (b)

Ramphele, M., & Wilson, F. (1989). "Uprooting Poverty: The South African Challenge." David Philip, Claremont.

Wittenberg, M. (2002). "Job Search in South Africa: A Nonparametric Analysis." South African Journal of Economics, 70(8), 1163-1196.

Wooldridge, J.M. 2003. 'Cluster-sample methods in applied econometrics',American Economic Review, p133-138, 2003, JSTOR

Zax, J. S. (1991). "Compensation for commutes in labor and housing markets." Journal of Urban Economics, 30(2), 192-207.

Zenou, Y. (2006). "Efficiency wages and unemployment in cities: the case of high-relocation costs." Regional Science and Urban Economics, 36(1), 49-71.