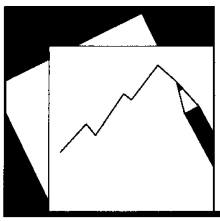
WP/15/xx



IMF Working Paper

The Costly Road to Work?
Wages and Transport Costs in South Africa

Maria Ngarachu, Axel Schimmelpfennig, Volker Schöer

PRELIMINARY DRAFT — NOT FOR QUOTATION

INTERNATIONAL MONETARY FUND

 WP/15/xx

IMF Working Paper

African Department

The Costly Road to Work? Wages and Transport Costs in South Africa

Prepared by Maria Ngarachu, Axel Schimmelpfennig, Volker Schöer

Authorized for distribution by [_]

[August 2015]

Abstract

Transport costs can potentially drive a wedge between the employer's wage costs and the worker's take-home pay. Transport costs can either push up the wage and thus unemployment or reduce living standards and contribute to inequality. In this paper, we investigate the relationship between transport costs and wages in South Africa using the National Income Dynamics Study. On average, higher transport costs are associated with higher wages, and we find some evidence for causality after accounting for possible endogeneity. The relationship differs across the income distribution, with workers at the bottom of the distribution finding it difficult to pass on transport costs in full. Indeed, transport costs can devour up to 24 percent of a worker's wage at the lower end of the income distribution, but only a small portion of the income at the higher end. As such, transport cost contribute to wage inequality.

JEL Classification Numbers: I31; J61; R23; R40

Keywords: Transport Costs; Wages; Instrumental Variables, South Africa

Author's E-Mail Address: Maria Ngarachu, maria@ngarachu.com (corresponding author); Axel Schimmelpfennig, aschimmelpfennig@imf.org; Volker Schöer, volker.schoer@wits.ac.za.

INTRODUCTION[1]

High unemployment, low labor force participation, and high inequality are three of the key challenges facing South Africa. The root causes of these challenges are multifaceted and interlinked. Zooming in on unemployment, factors considered to be important drivers in the literature include labor market rigidities related to regulations and their implementation, the wage bargaining framework, skill mismatches, and marginalization and segmentation in the labor market. Another factor that is often mentioned, but so far possibly under-researched, is spatial inequalities, an apartheid legacy. In short, a share of the population does not live close to economic centers, and the associated transport costs impede job search and drive a wedge between take-home wages and the cost to employers.

In this paper, we analyze this possible link between transport costs and aggregate labor market outcomes, namely the effect of transport cost on wages. Transport costs can impact labor market outcomes in a number of ways, in particular by pushing up wage demands or reducing search intensity. We use data from the National Income Dynamics Study (NIDS), a household panel, to shed light on the replationship between transport costs and wages in South Africa. For a job seeker, transport costs and the associated travel time could increase wage demands as individuals seek to recoup some of the pecuniary and opportunity costs of commuting. If job seekers have bargaining power, this would then raise the overall wage level, which could contribute to unemployment. If job seekers have no bargaining power at all, their take-home pay gets reduced by the incurred transport costs, which weighs on living standards and could contribute to inequality. Transport costs can also prevent the unemployed from looking for jobs in the first place, or reject a job offer if the wage is too low relative to the transport costs that have to be incurred.

Unpacking the relationship between transport costs and wages is potentially complicated by endogeneity and unobserved characteristics. It may be reasonable to assume that workers will try to recoup higher transport costs by negotiating higher wages, but it is equally plausable that higher wages allow workers to finance higher transport costs, i.e., allow them to move further away from their workplaces or use more expensive modes of transport. It is also possible that other simultaneous, unobserved factors drive the observed correlation between transport costs and wages. We control for potential endogeneity using the variation of petrol prices over time and across municipalities as an instrument for the reported transport costs. We estimate a Mincerian wage regression, including transport costs as a regressor in addition to the usual controls. Starting with a standard ordinary least squares, we test the robustness of our results using instrumental variable, Heckman selection and exploit the panel dimension of the NIDS data to control for individual fixed effects. Finally, we investigate if the relationship differs over the income distribution using quantile regressions.

Our findings show that higher transport costs are associated with higher wages, and that transport costs contribute to inequality. Even when controlling for endogeneity, we find a positive relationship between transport costs and wages, though the impact is smaller. The relationship between transport costs and wages differs across the income distribution, with

[1] We thank Neil Rankin for helpful comments on an earlier version of this paper.

workers at the bottom of the distribution finding it difficult to pass on transport costs in full, while workers at the top of the distribution find it easier. Our data also shows that transport costs can devour up to 24 percent of a worker's wage at the lower end of the income distribution, but only a small portion of the income at the higher end. As such, transport costs contribute to wage inequality, with the gini coefficient for the wage distribution increasing by 0.02 to 0.06 points.

The remaining paper is structured as follows: Section II provides a survey of the existing literature on transport costs and wages; Section III introduces the dataset and presents the descriptive statistics on the magnitude, distribution, and impact of transport costs across demographic groups; Section IV discusses the econometric methodology used and presents and the results; and, Section V concludes.

I. REVIEW OF THE LITERATURE

There is a wide academic literature seeking to explain unemployment in South Africa (see Fourie, 2011 for a survey). Key factors identified by this literature as explaining unemployment in South Africa include wage rigidities as a result of labor market regulation (Nattrass, 2000; Kingdon and Knight, 2006); centralized bargaining and the role of unions (Arora and Ricci, 2005); skills mismatch (Cichello, Fields, and Leibbrandt, 2005; Rodrik, 2006); and marginalization and segmentation in the labor market (see Fourie, 2011).

Akin to the above factors, is the notion that wages in the South African labor market are 'too high.' Accordingly, the South African literature focuses on the determinants of the reservation wages of the unemployed with specific emphasis on the role of unemployment duration, unemployment insurance, savings, wealth and social transfers such as the old age pension.[2] However, there is no consistent empirical evidence that high reservation wages are indeed responsible for unemployment in South Africa. Nattrass and Walker (2005) find that reservation wages are not out of sync with predicted wages. Similarly, Banerjee *et al* (2008) as well as Heintz and Posel (2008) conclude that excessive reservation wages cannot adequately account for the high levels of unemployment in South Africa. An exception to these findings is the study by Rankin and Roberts (2010) who differentiate between wages in different firm sizes. The authors find that even if reservation wages are similar to average predicted wages, they are above what young people could expect to earn in smaller firms.

One angle that is possibly less well researched is the role transport costs may have on wages. A starting point to understanding the impact of transport cost on wages is a basic search model. The reservation wage (w_r) is the optimal stopping point at which the income maximizing job seeker is perfectly indifferent between accepting the current wage offer and continuing to search for a better job offer.[3] Assuming that the job seeker is likely to have

[2] The reservation wage is the lowest wage at which an individual is willing to work. This concept is central to theoretical models of job search, labor supply and labor market participation (see, e.g., Mortensen, 1986, Maani and Studenmund, 1986; Holzer, Ihlanfeldt, and Sjoquist, 1994; Bloemen and Stancanelli, 2001; Prasad, 2003; Carolina and Pau, 2008; Lammers, 2009; Krueger and Mueller, 2011; and Krueger and Mueller, 2014).

[3] See for literature review: Rogerson, Shimer, and Wright, (2005), Search-Theoretic Models of the Labor Market: A Survey, Journal of Economic Literature, Vol. 43 (4): 959-988

access to non-labour income equal to b, faces search cost c_s, experiences a job offer arrival rate of λ, has a discount rate of r, and that jobs are exogenously separated at rate q, we can model the reservation wage of a representative job seeker in equation 1 as:

$$w_r = b - c_s + \frac{\lambda}{r+q}\int_{w_r}^{\infty}(w - w_r)dF(w) \quad (1)$$

Thus, we would expect the reservation wage of a job seeker to increase with non-labour income (b) and a higher job arrival rate (λ), while higher search cost (c_s), a larger discount rate (r) and a higher job separation rate (q) would affect the reservation wage negatively.[4] When comparing a wage offer to the reservation wage in this manner, the job seeker will focus on the take-home wage, excluding inter alia taxes and arguably transport costs. Thus, we could expect that higher transport costs would increase the reservation wage and subsequently the accepted wage once the job seeker transitions into wage employment.

Church, Frost and Sullivan (2000) as well as Rupert, Stancanelli and Wasmer (2009) note that transport costs are rarely addressed in the unemployment, search and wage literature. Where it has been addressed in the international literature (e.g., Rouwendal and Rietvield, 1994; van der Berg and Gorter, 1997; van Ommeren, 1998; Patacchini and Zenou, 2005; van Ommeren and Straaten, 2005; van Ommeren and Fosgerau, 2009; Verick, 2011; Phillips, 2011; Franklin, 2012) it is more often than not proxied by variables such as commuting distance and time, and studies predominantly focuses on its influence on search behaviour and/or subsequent employment outcomes (e.g., Holzer, 1994; Bartus, 2011; Sullivan, 2003). An exception is Laird (2006) who investigates the impact of commuting costs on the wage rates of the employed in Scotland. The author estimates three equations—the wage and commuting distance equations using two stage least squares whilst the commuting cost equation is populated using calibrated models—to investigate whether a relationship between wage rates and commuting costs exists. Thereafter, Laird (2006) sets out to determine whether lower commuting costs can be transmitted through the labour and transport markets into lower real wages and increased labour market participation. The research found evidence that wage compensation for commuting occurs, though only partially. Furthermore, the complex interaction between wage rates, commuting costs, work and household location decisions and the value of travel time means that, through labour supply effects, transport policy has little impact on wages.

The aforementioned use of proxies is mainly a result of the unavailability of data on actual commute costs. Nevertheless, Kain's (1968) seminal paper on the spatial mismatch hypothesis provides a good basis for the use of these proxies. The hypothesis seeks to explain the adverse labour market outcomes of minorities disconnected from job opportunities which is particularly relevant for South Africa given its history of segregation in the apartheid era. As such, physical accessibility from residential areas to job locations becomes important.

[4] Various empirical studies have confirmed these theoretical predictions (See, for instance, Lancaster and Chesher, 1983; Prasad, 2003; Brown and Taylor, 2008).

For South Africa, commentators frequently point to transport cost as a wedge between wage costs to the employer and take-home wages. Arora and Ricci (2005) and Mahajan (2011), for example, point out that transport costs are likely to drive up wages in the South African labor market sphere. Walker (2003) who investigates the determinants of reservation wages in Khayelitsha/ Mitchell's Plain (a metropolitan area in South Africa's Cape Town) hints at distance to work being important in the equation as individuals may be prepared to work for less if jobs are closer to home. Kingdon and Knight (2006) use indicators of remoteness (e.g., impassable roads and residence in a homeland), while Hinks (2008) uses distance from public transport to proxy commuting costs and show that these negatively affect the job searching process of the unemployed. Dimitrov (2010) looks at the accessibility of transport for low income individuals and those living in remote areas. Statistics South Africa (Stats SA) conducted a descriptive study, the National Household Travel Survey (NHTS), on behalf of the Department of Transport in 2003 and 2013 to better comprehend how and why South Africans travel. The survey found travel time followed closely by travel cost to be the most important factors influencing households' choice of transport.

To our knowledge, there is only one existing empirical study that makes uses of actual commuting costs in explaining wages in South Africa. De Lange (2013) explores the relationship between transport costs and reservation wages as well as the role of transport premiums (this refers to the premium added to wages required for travelling to job locations that are farther away from one's place of residence). The study uses the Labor Market Entry Survey (LMES) data set, comprising young Africans in three of South Africa's eight provinces. The author finds that individuals try to recoup a portion of their transport costs and that those who live closer to job centers recoup less. However, this study does not control for possible endogeneity.

Our paper attempts to add to the existing South African literature in three ways. First, we explicitly study the relationship—if any—of transport costs on wages, using actual data on transport costs for wage employed respondents from the NIDS. Second, we use an instrumental variable (IV) approach to account for a potential endogeneity bias. And third, we investigate if the potential effect from transport costs on wages varies across different income groups. We also provide detailed descriptive statistics of the incidence of transport costs across demographic groups and how transport cost impact inequality.

DATA AND DESCRIPTIVE STATISTICS

We make use of information on wages and transport costs contained in the first and second waves of the NIDS, a household survey that was fielded in 2008, 2010, and 2012 respectively.[5] Questions on wages and on transport costs with sufficient responses are only jointly available in the first and second waves. A total of over 28,000 individuals from approximately 7,300 households were interviewed in the first wave, with 18,864 individuals successfully interviewed in all three waves.

[5] A fourth wave is currently being conducted with the data set to be released in 2016.

For all descriptive statistics and regressions analysis, we use weights provided in the NIDS to achieve nationally representative results. The original wave 1 design weights ensure that the age-sex-race marginal totals in NIDS match that of the population. However, using these in the second wave would not be entirely accurate as they would not account for the household and individual level attrition between the first two waves. Thus, we use post-stratification calibrated weights[6] in our cross-sectional Wave 2 analyses.

Similarly, individuals who were successfully re-interviewed in the second wave are not a random subset of all individuals surveyed in the first wave. Accordingly, panel weights—the product of the probability of being interviewed in the first wave, multiplied by the probability of being successfully re-interviewed in the second wave, conditional on appearing in the first wave—are used in the panel to account for this non-random attrition.

A. Commuting to Work—Descriptive Statistics

One impression in the literature on the South African labor market is that many workers have to travel long distances in their daily commute. In our data, the average daily one-way commute is 19 km (Table 1 and Figure 1).[7] There is some variation in the average and maximum distances travelled across the different demographic groups, but looking at the kernel densities (or the median), commuting distances appear fairly similar across these groups. In general, such distances are not necessarily out of line with what we know anecdotally from other countries.

Table 1: One-way Commuting Distance by Race, 2010 (Kilometers per Day)

	Race				
	All	African/ Black	Colored	Asian/ Indian	White
Median	10	10	10	15	11
Mean	19	19	17	25	20
Standard Deviation	25.74	27.03	21.36	23.69	23.50
Minimum	0	0	0	0	0
Maximum	275	275	200	87	245
Observations	3,178	2,229	701	48	200

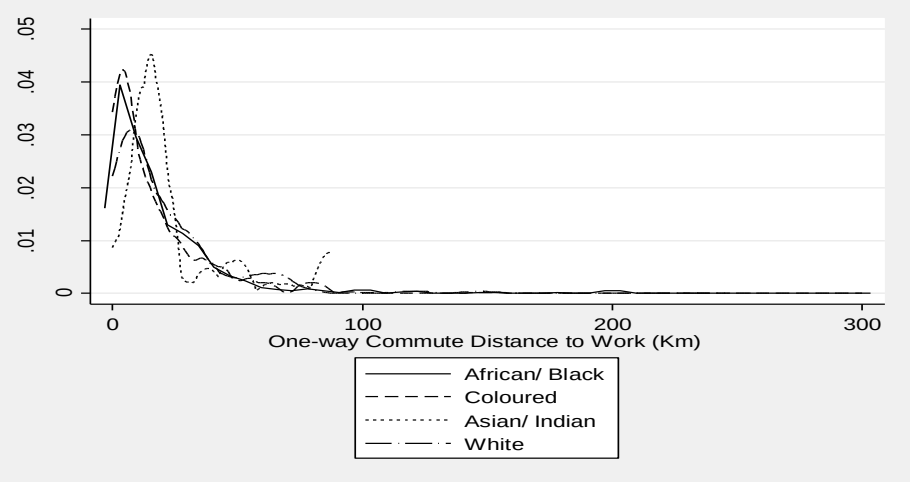
Source: NIDS Wave 2. Authors' calculations using design weights.[8]

[6] For a detailed description on the post-stratification weights, see Wittenberg (2009)

[7] Distance and time questions pertaining to work were not asked in the first wave

[8] On occasion, post-stratification/ sampling weights yield very slightly different means. However, making use of them does not permit a more detailed distribution of this mean.

Figure 1: Density Distribution of One-way Commuting Distance by Race, 2010 (Kilometers per Day)



Source: NIDS Wave 2. Authors' calculations using design weights.

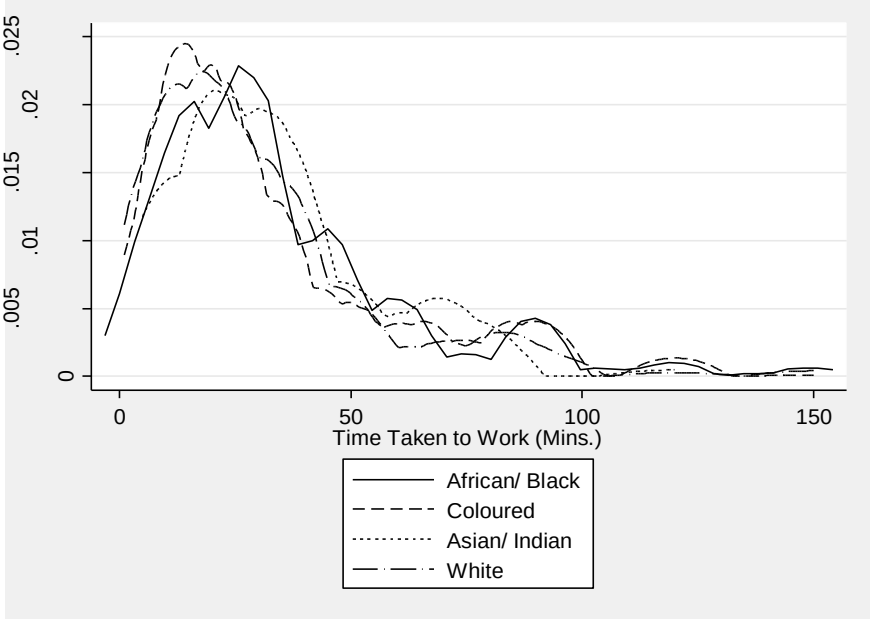
Even short distances can be associated with long commute times, if there is no direct and convenient means of transport available. In our sample, the median daily one-way commute time is 30 minutes and there is not much variation across the demographic groups (Table 2 and Figure 2). Again, these magnitudes do not seem out of line with anecdotal impressions from other countries (see for example Bartus, 2011, where Dutch workers spend on average 40 minutes on the road).

Table 2: One-way Commute Time by Race, 2010 (Minutes per Day)

	Race				
	All	African/ Black	Colored	Asian/ Indian	White
Median	30	30	25	30	25
Mean	34	36	33	33	30
Standard Deviation	26.73	27.12	27.30	21.92	25.08
Minimum	3	3	3	5	3
Maximum	150	150	150	120	150
Observations	3,599	2,600	745	50	204

Source: NIDS Wave 2. Authors' calculations using design weights.

Figure 2: Density Distribution of One-Way Commute by Race, 2010 (Minutes per Day)



Source: NIDS Wave 2. Authors' calculations using design weights.

Lastly, it is transport costs in rand that are likely to matter most when thinking about wages. The relevant question in the NIDS is: "*How much did you spend last month on transport to and from this job?"* The question is posed only to the wage-employed and not to the self-employed.[9] Thus, we focus only on the wage-employed in this paper, and the "employed" refers to this group of individuals. For these, the average transport costs are rand 342 and rand 497 per month in 2008 and 2010 respectively (Table 3). There are some interesting differences across demographic groups. Men report higher transport costs than women.[10] And, transport costs seem highest for Whites and Indians (Table 4). Transport costs also vary substantially. Some respondents report zero transport costs which could either imply that they walk to work or this might be misreporting. At the other end, transport costs as high as rand 25,000 which is likely to reflect misreporting and may include a weekly or monthly commute away from one's primary residence closer to work rather than the daily commute.

[9] The NIDS also has information on transport costs for the unemployed. Although one could predict a wage for these individuals based on their characteristics, the question on transport costs faced by the unemployed - the majority of the unemployed reported an expenditure of "zero" on work-related transport - seems less clear and more relevant for search intensity.

[10] The difference in 2008 is not statistically significant

Table 3: Transport Costs by Gender (rand[1/] per Month)

	Gender[2/]					
	Male		**Female**		**Total**	
	2008	2010	2008	2010	2008	2010
Median	205	392	222	360	222	380
Mean	332	539	355	447	342	497
Std. Dev.	424	662	410	412	419	565
Minimum	0	0	0	0	0	0
Maximum	3,003	3,500	3,114	3,200	3,114	3,500
Observations	1,820	1,256	1,612	1,303	3,432	2,559

Source: NIDS Waves 1 and 2. Authors' calculations using cross-sectional design weights.

1/ Expressed in constant 2010 prices.
2/ Although the average monthly transport cost of men appears to be considerably higher than that of women; two tailed t-tests indicate that there aren't significant differences between the two groups.

Table 4: Transport Cost by Race (rand[1/] per Month)

	African/ Black		**Colored**		**Indian**		**White**	
	2008	2010	2008	2010	2008	2010	2008	2010
Median	167	300	133	300	779	600	556	600
Mean	267	391	309	425	795	805	673	873
Std. Dev.	334	406	399	538	507	799	605	801
Min.	0	0	0	0	0	0	0	0
Max.	3,114	3,500	2,669	3,500	2,224	3,000	3,003	3,500
Obs.	2,348	1,909	724	421	58	46	293	182

Source: NIDS Waves 1 and 2. Authors' calculations using cross-sectional design weights.

1/ Expressed in constant 2010 prices.
2/ Weighted two tailed t-tests indicate that there are significant differences between the transport costs of Africans and Coloreds and that of Indians and Whites.

Decomposing transport costs by occupation, one can see that individuals in the upper occupation categories incur greater costs than those in the lower categories (Table 6). This comes as no surprise as they should have greater flexibility in choosing their preferred place of residence and mode of transport. Thus, for the higher skill occupations, the direction of causality might run from the wage level to transport cost rather than transport cost affecting wages. Delving further into specific occupations in which we would expect individuals to incur little to no daily commute costs, one can confirm from Table 7 that this is indeed the case—with the exception perhaps for mineworkers.[11] In 2008, farmworkers in the dataset did

[11] In South Africa, farmworkers, mineworkers, and domestic workers often live at their place of work during the week or an extended period of time, and return to their primary residence (and their families) only in intervals.

not incur any transport costs whereas domestic workers incurred costs as little as rand 56 per month.

Table 5: Transport Cost by Occupation (rand[1/] per Month)

	Median		Mean		Max.[2/]		Obs.	
Occupation	2008	2010	2008	2010	2008	2010	2008	2010
Armed forces	-	510	-	567	-	960	0	7
Managers	667	800	826	1,106	2,781	3,500	122	96
Professionals	345	500	524	695	3,114	3,500	415	353
Technicians and assoc. professionals	445	550	411	795	2,224	3,000	131	131
Clerical support	417	440	496	463	2,669	2,100	283	137
Service and sales	222	320	267	471	2,447	3,500	431	429
Skilled agricultural, forestry and fishery	0	250	33	333	723	600	356	17
Craft and related trades	222	300	322	345	3,003	2,500	444	242
Plant and machinery	133	300	255	337	2,558	2,000	314	274
Elementary occupations	17	240	153	266	1,780	2,000	877	721

Source: NIDS Waves 1 and 2. Authors' calculations using cross-sectional design weights.

1/ Expressed in constant 2010 prices.

2/ Due to space constraints, we omit the minimum reported transport cost – with the exception of a R250 minimum in the armed forces category in 2010, the minimum for all other categories in both years is R0.

Table 6: Transport Cost by Occupation (rand[1/] per Month)

	Median		Mean		Max.[2/]		Obs.	
Occupation	2008	2010	2008	2010	2008	2010	2008	2010
Farmworkers [3/]	0	350	28	297	723	2,000	560	75
Mineworkers [4/]	234	480	413	397	2,224	1,110	89	38
Domestic workers	56	-	142	-	890	-	348	0

Source: NIDS Waves 1 and 2. Authors' calculations using cross-sectional design weights.

1/ Expressed in constant 2010 prices. 2/ Due to space constraints, we omit the minimum reported transport cost – with the exception of a R250 minimum in the armed forces category in 2010, the minimum for all other categories in both years is R0. 3/ Includes all those in the skilled agricultural, forestry and fishery occupation category as well as motorized farm and plant operators (plant and machinery category); and, those in elementary positions. 4/ Includes all mining, mineral and metal processing operators as well as metal, cement and other mineral products machine operators (plant and machinery category); and, mining and constructional labourers (elementary category)

Another interesting dimension is the extent to which commute costs differ among the different types of settlements. Although the figures in 2008 are more in line with what one would expect (on the lower end, zero rand among those in rural formal settlements; on the upper end, rand 278 among those residing in urban formal settlements), the figures in 2010 are slightly more variant (Table 7). Furthermore, looking at both years simultaneously, those in urban informal settlements seem to experience relatively similar commute costs with those residing in urban formal areas. This is in spite of a considerable proportion of Africans residing in urban informal areas whom we saw experience much lower transport costs than Whites who dominate urban formal establishments (Table 8). Having said that, this peculiar result may be because it is much easier for Africans residing in urban informal areas to compute their monthly transport costs—50 percent use taxis—compared to those in urban formal areas—37 percent use motor vehicles—who may only state the fuel consumption and not incorporate the full wear and tear of the vehicle (Table 9).

Table 7: Transport Cost by Region (rand[1/] per Month)

	Region							
	Rural formal		**Tribal authority areas**		**Urban formal**		**Urban informal**	
	2008	2010	2008	2010	2008	2010	2008	2010
Median	0	200	111	300	278	400	245	280
Mean	124	350	211	379	420	547	260	320
Std. Dev.	296	536	265	363	463	560	220	422
Min.	0	0	0	0	0	0	0	0
Max.	2,781	3,000	1,713	3,000	3,114	3,500	2,224	3,500
Obs.	714	284	593	510	1,902	1,478	223	253

Source: NIDS Waves 1 and 2. Authors' calculations using cross-sectional design weights.

1/ Expressed in constant 2010 prices.

Table 8: Race by Region (Percentage)

	Region							
	Rural formal		**Tribal authority areas**		**Urban formal**		**Urban informal**	
	2008	2010	2008	2010	2008	2010	2008	2010
African/ Black	78.30	69.18	100.00	99.80	64.55	56.46	95.83	97.39
Colored	11.52	13.08	-	0.20	12.17	13.72	4.17	2.61
Indian	7.95	15.26	-	-	3.14	4.08	-	-
White	2.23	2.48	-	-	20.14	25.74	-	-
Obs.	711	284	592	510	1,897	1,477	223	253

Source: NIDS Waves 1 and 2. Authors' calculations using cross-sectional design weights.

Table 9: Transport Mode by Region, 2010[1/] (Percentage)

	Region			
	Rural formal	Tribal authority areas	Urban formal	Urban informal
Train	0.9	1.2	5.7	12.9
Bus	4.5	19.2	6.2	8.4
Taxi	13.0	27.3	25.0	50.9
Motor vehicle	20.8	12.0	37.1	4.4
Bicycle	1.5	0.5	0.8	4.6
On foot	36.7	25.4	15.6	8.1
Employer provided transport	16.1	11.8	8.2	9.4
Work from home	3.1	2.2	1.0	1.2
Other	3.5	0.4	0.4	0.1
Observations 2/	630	813	2,060	316

Source: NIDS Waves 1 and 2. Authors' calculations using cross-sectional design weights.

1/ The survey in 2008 did not include a question on the mode of transport to work.
2/ All transport mode reporters – respondents' who only selected either "Bicycle" or "On foot" were not asked about their transportation costs.

Looking at transport costs in absolute terms may be misleading. As alluded to earlier, some individuals may happily pay a high price for transport if they can afford it and if it suits their preferences say for expensive cars or particular (remote) residential neighborhoods. Others may see their budgets severely impacted if transport costs are high compared to their wages. Ultimately, it is the transport costs relative to the wage that will matter from a social perspective: in how far do transport costs lower a worker's take-home pay. In our sample, the difference in the transport cost to wage ratio across income groups is striking, though not unexpected (Table 10). For the lowest income group, transport costs account for up to 24 percent of take-home pay. In contrast, the impact of transport costs on the highest income group (whose transport costs as a share of income are approximately three times less than that of the poorest) is considerably less severe.

Table 10: Proportion of Transport Costs to Actual Wages by Income Decile, 2010[12]

Percentile 1/	10th	20th	30th	40th	Median	60th	70th	80th	90th
Transport cost 1/	305	404	402	555	455	490	490	749	778
Actual wage 2/	1,831	1,657	3,062	3,321	3,702	4,770	6,198	10,292	10,421
Share of income 3/	0.17	0.24	0.13	0.17	0.12	0.10	0.08	0.07	0.07

Source: NIDS Wave 2. Authors' calculations using design weights.

1/ Average transport cost and actual wage of individuals in the given income decile. With the exception of a marked increase to 29% of income for the second lowest income category, making use of the median transport cost and actual wage leaves our proportions relatively unchanged.
2/ We exclude the 7 individuals who report an actual wage of "zero".
3/ Transport cost/Actual wage

The distributional impact of transport costs can be summarized with the gini coefficent. We compare the gini coefficient for the reported wages and the gini coefficient for the reported wages less reported transport cost—a "take-home" pay before taxes (Table 11). The gini for reported wages is 0.55 in 2008 and 0.44 in 2010.[13] The decline is remarkable, and may have to do with the large retrechments (one million jobs lost) that occurred during the period, with retrenchments typically falling mainly on low-wage workers. When looking at the "take-home" pay, after transport costs, the gini coefficient increases by 0.02 in 2008 and 0.06 in 2010, illustrating the notable impact of transport costs on inequality.

Table 11: Gini Index

	2008	**2010**
Reported wage	0.55	0.44
Take-home wage (wage minus transport costs)	0.57	0.50

Source: NIDS Waves 1 and 2. Authors' calculations.

B. Wages—Descriptive Statistics

The pattern of reported wages in our sample across demographic groups is consistent with the literature (e.g., Kingdon and Knight, 1999). Men report higher wages than women (Table 12). Whites and Indians have significantly higher wages than Africans and Coloreds (Table 13). The latter could be partly explained by differences in skill levels that trace back to Africans and Coloreds having been previously excluded from equal education and economic opportunities.

[12] We present the second wave results as this had more observations - the 2008 proportions are relatively similar

[13] The estimated gini coefficients in this paper differ from the number typically cited for South Africa—a gini of about 0.77. The difference arises because the gini here is calculated for individual wage earnings whereas the gini typically looked at is calculated for household income.

Table 12: Actual Wages by Gender (rand[1/] per Month)

	Gender					
	Male		**Female**		**Total**	
	2008	2010	2008	2010	2008	2010
Median	2,781	3,500	1,891	2,350	2,224	3,000
Mean	4,263	5,224	3,675	4,075	4,015	4,721
Std. Dev.	4,157	6,081	3,523	4,389	3,910	5,436
Min.	167	4	167	10	167	4
Max.	25,581	37,000	24,469	32,000	25,581	37,000
Obs.	1,616	1,754	1,422	1,738	3,038	3,492

Source: NIDS Waves 1 and 2. Authors' calculations using cross-sectional design weights.

1/ Expressed in constant 2010 prices.
2/ With the inclusion of outliers, the three highest monthly wages are R 70,000, R 75,000 and R 90,000

Table 13: Actual Wages by Race (rand[1/] per Month)

	African/ Black		**Colored**		**Indian**		**White**	
	2008	2010	2008	2010	2008	2010	2008	2010
Median	2,002	2,400	2,753	3,100	4,004	9,300	7,786	8,000
Mean	3,042	3,365	4,258	4,828	6,375	10,722	8,701	11,102
Std. Dev.	2,961	3,244	3,689	4,966	5,653	8,103	5,124	8,756
Min.	167	4	133	12	1,023	1,000	222	500
Max.	25,581	33,000	22,245	32,000	24,469	31,000	25,581	37,000
Obs.	2,103	2,556	627	725	40	41	259	168

Source: NIDS Waves 1 and 2. Authors' calculations using cross-sectional design weights.

1/ Expressed in constant 2010 prices

We also see wages increasing with age up to a point (Table 14). This is to be expected as individuals accumulate experience over time. However, consistent with search theory, for the oldest age brackets, there is some evidence of wages coming down again as the return to finding the perfect match decreases with fewer years remaining in the economically active phase of one's life.

Table 14: Actual Wages by Age Group (rand[1/] per Month)

	Median		Mean		Max.[2/]		Obs.	
Age Group	2008	2010	2008	2010	2008	2010	2008	2010
<18	1,112	2,000	817	1,541	1,212	2,000	8	3
18-29	2,002	2,200	2,917	3,304	17,796	25,000	746	917
30-45	2,669	3,000	4,161	4,566	25,581	33,000	1,452	1,600
46-65	2,669	4,000	4,871	6,486	25,581	37,000	812	952
>65	1,335	1,868	2,643	3,246	20,020	20,000	17	20

Source: NIDS Waves 1 and 2. Authors' calculations using cross-sectional design weights.

1/ Expressed in constant 2010 prices
2/ The minimum wage in 2008 is R300 for the over 65's and R150 for the other age groups. In 2010, the minimum wage is R300 for the under 18's; R50 for the over 65's; and ranges between R4 and R16 for the other three categories.

Wages also differ by where people live. Individuals in urban formal areas receive substantially more than those in other areas (Table 15). Of interest to note is that individuals residing in urban informal settlements have the lowest reported wages—lower than those of individuals in rural formal and tribal authority areas. Moreover, in spite of individuals in these settlements facing similar commute costs to those in other settlements—particularly urban informal (see Table 7)—the reported wages are quite disparate.

Table 15: Actual Wages by Region (rand[1/] per Month)

	Region							
	Rural formal		**Tribal authority areas**		**Urban formal**		**Urban informal**	
	2,008	2,010	2,008	2,010	2,008	2,010	2,008	2,010
Median	1,212	1,450	1,446	1,800	3,448	4,000	1,446	2,000
Mean	1,738	2,415	2,272	2,822	5,097	5,960	1,820	2,276
Std. Dev.	1,514	2,746	2,008	2,846	4,359	6,230	1,518	1,642
Min.	167	12	167	20	167	4	222	120
Max.	18,908	19,000	14,230	25,000	25,581	37,000	20,020	20,000
Obs.	660	653	510	717	1,670	1,793	198	277

Source: NIDS Waves 1 and 2. Authors' calculations, using cross-sectional design weights.

1/ Expressed in constant 2010 prices

ECONOMETRIC APPROACH AND RESULTS

A. Econometric Approach

Following the standard Mincerian wage regression, the basic equation to test whether wages are influenced by transport costs is:

$$Y_{i} = \alpha_i + \beta_0 Tr_i + \beta_1 X_{1i}' + \beta_1 X_{2i}' + \beta_1 X_{3i}' + \beta_1 X_{4i}' + \varepsilon_i \quad (2)$$

where:

Y_i is an individual's self-reported monthly wage; and

Tr_i is the self-reported monthly transport cost to an individual's work place.

In addition, we control for demographic, household, employment and regional covariates:

> X_{1i}' is a vector of individual-specific demographics and includes age, gender, marital status, race and education;
>
> X_{2i}' is a vector of controls for household composition and includes household income per capita and a dummy variable to indicate whether an individual is the main breadwinner;
>
> X_{3i}'controls for employment characteristics, specifically, occupation category and industry;
>
> X_{4i}' consists of regional indicators which include the regional location and province of the household;

ε_i is the disturbance term.

We estimate the basic equation in a log form, using different econometric techniques. The log-log specification gives the elasticity of wages with respect to transport costs whereas a linear specification would give the wage increase for a unit increase in transport costs.[14] The model is estimated, using a number of different econometric approaches to address possible estimation issues:

Ordinary Least Squares (OLS) for the 2008 and 2010 waves individually and pooled;

Instrumental Variables (IV), using district-specific petrol prices as an instrument for transport costs, again for the 2008 and 2010 waves individually and pooled;

Heckman selection into employment with IV;

Fixed-effects panel with and without IV;

[14] The literature uses mostly the log-log specification, though some papers (see for example, Laird, 2006, who tests several forms of the model) use the linear specification.

Quantile regressions to allow for different relationships between transport costs and wages across the income distribution.

The choice of covariates follows the literature though. Age, gender, marital status, race and education improve the performance of the model.[15] Although the indicator variable for whether or not an individual is the main breadwinner in the household is individually insignificant, it enriches the explanatory power of the model and hence it is kept in together with the total income per capita of the household.[16] Our exclusion of the potential experience and its squared equivalent from our employment characteristics is due to the fact that albeit enhancing the model, they are individually insignificant and are incorrectly signed which is to be expected given the high correlation with some of the other covariates.[17] They also render the higher education categories and all industries insignificant. Lastly, although there were slight changes in province demarcation from the 2001 to the 2011 Census, we use the 2001 demarcations in the regional indicators as these are the ones which prevailed when the second wave was conducted.

As usual in household surveys, our sample includes outliers as easily seen from the descriptive statistics. In line with the literature, we exclude outliers for the key variables of interest from the sample. For wages, this entails excluding the top and bottom 1 percent. For transport costs, however, the self-reported cost only changes from "zero" to a positive value at the 13th percentile. Hence, we cannot exclude the bottom 1 percent of reported transport costs. Instead, we make use of a log transformation that can deal with zeros—the inverse hyperbolic sine transformation: log (x+(x2+1)1/2), and exclude only the top 1 percent of transport costs. However, there is a question whether zero reported transport costs are reporting errors, while it is conceivable that some may indeed perceive zero transport costs if they are walking to work, biking to work, or living at their work place. We report regression results for a sample including and excluding observations with zero transport costs.

Regression Results

The main results are summarized in Tables 16 through 22. The corresponding detailed regression results are shown in Tables A1 through A11 in the appendix.

[15] Age squared, an interactive term between gender and marital status, and an indicator variable for disability did not enhance the model. Moreover, due to the high correlation between these variables and some of the other variables in the model; in some instances, we had insignificance of variables which the literature indicates should be significant. We exclude the number of children from our model as including them considerably reduces our sample size. However, we do include household size which partly encompasses this.

[16] The components of aggregate household income are labor market income, government grants, other government income, investment income, remittances received, subsistence agriculture income and imputed rent for owner-occupied housing.

[17] As with most of the literature, this is computed as age - number of years of education - 6

Ordinary Least Squares

Looking at the individual wave OLS specifications, we find a positive relationship between transport costs and wages. Although the magnitude is considerably higher with the exclusion of zero transport cost reporters (columns 1b), we see that a 1 percent increase in costs of transport is associated with a 0.05 and 0.02 percent increase in wages in 2008 and 2010 respectively. The control variables are mostly significant and carry the expected signs.[18] Combining the two waves confirms this result giving an elasticity of 0.04.

Instrumental Variables

Transport costs could be correlated with the error term if there are unobserved characteristics that influence the location of residence and the wage. As a simple example, some individuals may be highly motivated which causes them to move closer to economic centers and at the same time increase their earnings potential. In this case, OLS coefficients would be biased, and we have to use IV.

To test for this potential endogeniety, we instrument for tranport costs using monthly petrol prices.[19] Petrol prices should be a major factor influencing transport costs. At the same time, they are exogenous to the individual's decision where to live. In South Africa, retail petrol prices are set monthly at a district level, based on a formula that, boadly speaking, accounts for the import price, the cost of transporting the fuel to the point of sale (the district), and allows for a profit margin. We can match the petrol price to an individual in our sample for the month in which the interview took place and the district that the individual lives in.

The IV regression results for the individual waves and the pooled sample support the OLS results and show that higher transport costs increase wages. Consistent with the usual findings, the estimated transport cost parameter becomes smaller (in one instance negative), and "loses" significance.

Selection into Employment (Heckman Selection)

Lastly, it is important to account for selection into employment. For the first stage of the Heckman selection, we make use of the number of children as well as whether an individual's health status is poor as the exclusion restrictions. Although there is evidence of

[18] The literature on economic geography suggests that rental costs for housing could also matter for wages (White, 1986; Van Ommeren, 1996; Smith and Zenou, 2003). Living closer to city centers is often associated with higher rents for similar types of dwellings, and this may drive up wages, while transport costs should decline, with an offsetting impact on the wage. We can study this relationship, but it reduces our sample further. In line with the hypothesis, preliminary results (not shown) suggest that the rental variable is highly significant. Even so, although the magnitude of the elasticity of wages with respect to transport costs is slightly reduced; it remains highly significant. The other covariates are largely unaffected by including rent in the list of regressors.

[19] We used the Hausman-Wu test to test for the endogeneity of transport costs; the Sargan statistic for the validity of additional instruments; and, the Stock and Yogo test for the identification of weak instruments.

selection in 2010, the overall result is unchanged. Pooling the two waves, we see that selection effects into employment are negligible.

Table 16a: Effect of Transport Costs on Wages, 2008[20]

Log wage	OLS		IV		HECK + IV	
	1a)	1b)	2a)	2b)	3a)	3b)
Log transport cost	0.0459***	0.184***	-0.140*	0.675*	-0.245	0.453*
	(0.0061)	(0.0303)	(0.0735)	(0.346)	(0.203)	(0.245)
Inverse mills ratio					-0.858	-0.153
					(0.533)	(0.343)
Constant	6.429***	5.793***	6.784***	3.11	8.513***	4.364***
	(0.163)	(0.254)	(0.214)	(1.926)	(1.082)	(1.646)
Observations 1/	2,481	1,233	2,328	1,149	997	548
R-squared	0.671	0.682	0.422	0.536	0.255	0.71

Notes: Standard errors are clustered at the household level and sampling weights are included. Strata with single sampling unit are centered at the overall mean. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

1/ Though we report the results for the maximum sample for each specification, we can confirm that our results remain relatively unchanged for the sample of 997 when we include zero transport cost reporters and 548 when we exclude them

[20] The a)'s represent our main regression with zero transport costs included while the b)'s represent the regressions where zero's are considered outliers and excluded from the model

Table 16b: Effect of Transport Costs on Wages, 2010

Log wage	OLS		IV		HECK + IV	
	1a)	1b)	2a)	2b)	3a)	3b)
Log transport cost	0.0263***	0.181***	0.411*	1.523	0.248	0.836
	(0.0114)	(0.0344)	(0.231)	(1.924)	(0.261)	(1.132)
Inverse mills ratio					-1.720**	-1.306**
					(0.780)	(0.511)
Constant	6.431***	5.673***	4.922***	-0.951	7.766***	3.446
	(0.015)	(0.302)	(0.943)	(9.452)	(0.781)	(6.588)
Observations 1/	1,894	1,496	1,799	1,417	826	689
R-squared	0.578	0.596	0.010	-0.233[21]	0.534	0.485

Notes: Standard errors are clustered at the household level and sampling weights are included. Strata with single sampling unit are centered at the overall mean. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

1/ Though we report the results for the maximum sample for each specification, we can confirm that our results remain relatively unchanged for the sample of 826 when we include zero transport cost reporters and 689 when we exclude them

Table 17: Effect of Transport Costs on Wages, Pooled (2008 and 2010)

Log wage	OLS		IV		HECK + IV	
	1a)	1b)	2a)	2b)	3a)	3b)
Log transport cost	0.0354***	0.199***	0.120	0.812*		
	(0.0047)	(0.091)	(0.0961)	(0.523)		
Inverse mills ratio						
Constant	6.200***	5.663***	6.009***	2.41		
	(0.465)	(0.268)	(0.262)	(0.224)		
Observations	4,375	2,729	4,127	2,566		
R-squared	0.622	0.575	0.317	0.296		

Notes: Standard errors are clustered at the household level and sampling weights are included. Strata with single sampling unit are centered at the overall mean. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

[21] Mathematically, a negative R squared indicates a poor fit of the data: the weak instrument in this regression can confirm this. However, we are also cognizant of the R squared being limited as a measure of goodness of fit in an instrumental variable framework particularly in light of the fact that the model was significant.

Table 18: Effect of Transport Costs on Wages, Panel (2008 and 2010)

Log wage	FE		FE + IV	
	1a)	1b)	2a)	2b)
Log transport cost	0.0256**	0.102**		
	(0.0128)	(0.0329)		
Constant	6.171***	5.624***		
	(0.277)	(0.142)		
Observations	2,312	1,197		

Notes: Individual fixed effects and panel weights are used. Strata with single sampling unit are centered at the overall mean. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Panel

<< insert once we receive results from data center >>

Quantile Regression Approach

As the burden of transport cost falls differently on the various income levels, we explore whether the relationship between transport costs and wages differs across income groups. Standard linear regression techniques summarize the average relationship between a set of regressors and the outcome variable based on the conditional mean function E(y|x). As discussed above, there are reasons to at least consider that the relationship between transport costs differs across the distribution of wages. At the lower end of the distribution, it may be more difficult to pass on transport costs to wages than it is at the upper end. Quantile regressions, first introduced by Koenker and Bassett (1978), allow us to describe the relationship at different points in the conditional distribution of y, and thus analyze whether the ability to pass on transport cost is different for different levels of the wage.

If we consider the standard linear regression model with intercept, a, and k x 1 slopes **b**;

$$y = a + \mathbf{b}\mathbf{x} + u$$

our previous OLS parameters were estimated by solving the following equation:

$$\min_{a,\mathbf{b}} \sum_{i=1}^{N} (y_i - a - \mathbf{x}_i\mathbf{b})^2$$

Correspondingly, the quantile regression is described by:

$$Quant_{\tau}(\mathrm{y_i} \mid \mathbf{x_i}) = \alpha(\tau) + \mathbf{x_i}\boldsymbol{\beta}(\tau$$

where the τ^{th} $(0 < \tau < 1)$ quantile regression estimator β_τ is solved by minimizing asymmetrically weighted sum of absolute standard errors:

$$Q(\beta_\tau) = \sum_{i:y_i \geq x_i'\beta}^{N} \tau \mid y_i - \alpha_\tau - \mathrm{x'}_i \beta_\tau \mid \; + \sum_{i:y_i < x_i'\beta}^{N} (1-\tau) \mid y_i - \alpha_\tau - \mathrm{x'}_i \beta_\tau \mid$$

Quantile regressions thus provide a richer characterization of the data permitting us to consider the impact of transport costs on the entire distribution of reservation wages and not merely on its conditional mean.[22] Quantile regressions are also more robust to outliers — akin to the median being more robust compared to the average — which as we saw, influence our results. Moreover, they are robust to non-normal errors which could be the case for wages. Due to the nature of the regression technique, we retain outliers unlike in the preceding specifications.

The results show that the transport coefficient is significant, especially for the lower income groups. Importantly, the estimated coefficient increases when zero transport cost reporters are excluded from the sample. What does this imply? Looking at the 2010 wave estimated for observations with zero transport costs excluded, we find that individuals can recoup higher transport costs, and much more so in the upper income deciles: At the 10^{th} decile, a rand 12.76 increase in transport costs is associated with a rand 20.94 increase in the wage, whereas for an individual in the 90^{th} income decile, a rand 44.52 increase in transport costs is associated with a wage increase of rand 138.60. Thus, at the lower end of the income distribution, wages more or less keep up with transport costs, while individuals in the higher income groups seem to be able to overcompensate, a finding that would be consistent with the endogeneity view: higher wages allow individuals to choose more expensive transport modalities. However, the pictures changes substantially when looking at the sample with zero transport costs. Now, no income group is able to recoup higher transport costs through higher wages. Looking again at the 2010 wave, at the 10th decile, a rand 18.30 increase in transport costs is associated with a rand 2.24 increase in the wage, whereas for an individual in the 90th income decile, a rand 46.68 increase in transport costs is associated with a wage increase of rand 20.70. Thus, at the lower end, workers can only recoup around 12% of the increase in the transport costs while the upper end can recoup 44%. These different results warrant further investigation on who the individuals who report zero transport costs are.

<< insert quantile IV regressions once this is feasible at data center >>

[22] The same exercise was conducted for each of the demographic groups separately (not shown), but no meaningful results were obtained.

Table 19a: Effect of Transport Costs on Wages by Income Decile, 2008

Percentile	10th	20th	30th	40th	Median	60th	70th	80th	90th
Wage 1/ 2/	1,096	1,589	2,462	2,675	3,089	3,590	5,572	7,502	9,077
Transport cost 2/	213	230	300	274	469	476	474	745	742
Proportion of income 3/	0.19	0.14	0.12	0.10	0.15	0.13	0.09	0.10	0.08
Elasticity 4/	0.0390***	0.0438***	0.0457***	0.0380***	0.0370***	0.0322***	0.0277***	0.0261***	0.0217***
Marginal effect:									
6% increase in transport cost	12.76	13.77	18.00	16.47	28.14	28.56	28.47	44.68	44.52
% increase in wages	0.23	0.26	0.27	0.23	0.22	0.19	0.17	0.16	0.13
Rand increase in wages	2.34	3.94	5.48	6.84	7.77	9.66	9.97	14.09	19.53

Source: NIDS Wave12. Authors' calculations – design weights used

1/ Individuals who report a wage of "zero" are excluded from the analysis
2/ Average wage and transport cost of individuals in the given income decile
3/ Transport cost/Wage
4/ Bootstrapped standard errors used. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 19b: Effect of Transport Costs on Wages by Income Decile, 2010

Percentile	10th	20th	30th	40th	Median	60th	70th	80th	90th
Wage 1/ 2/	1,831	1,657	3,062	3,321	3,702	4,770	6,198	10,292	10,421
Transport cost 2/	305	404	402	555	455	490	490	749	778
Proportion of income 3/	0.17	0.24	0.13	0.17	0.12	0.10	0.08	0.07	0.07
Elasticity 3/	0.0374**	0.0462***	0.0356***	0.0396***	0.0313***	0.0263***	0.0221***	0.0220***	0.0230***
Marginal effect:									
6% increase in transport cost	18.30	24.24	24.12	33.30	27.30	29.40	29.40	44.94	46.68
% increase in wages	0.22	0.28	0.21	0.24	0.19	0.16	0.13	0.13	0.14
Rand increase in wages	2.24	4.16	4.27	7.13	6.57	7.89	7.96	11.88	20.70

Source: NIDS Wave 2. Authors' calculations – design weights used

1/ Individuals who report a wage of "zero" are excluded from the analysis
2/ Average wage and transport cost of individuals in the given income decile
3/ Transport cost/Wage
4/ Bootstrapped standard errors used. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 20a: Effect of Transport Costs on Wages by Income Decile, 2008 (Zero transport cost reporters excluded)

Percentile	10th	20th	30th	40th	Median	60th	70th	80th	90th
Wage 1/ 3/	1,096	1,589	2,462	2,675	3,089	3,590	5,572	7,502	9,077
Transport cost 3/	198	212	239	254	369	453	476	623	723
Proportion of income 4/	0.18	0.13	0.10	0.10	0.12	0.13	0.09	0.08	0.08
Elasticity 5/	0.198***	0.199***	0.199***	0.182***	0.177***	0.187***	0.180***	0.160***	0.178***
Marginal effect:									
6% increase in transport cost	11.88	12.72	14.34	15.27	22.14	27.18	28.59	37.36	43.38
% increase in wages	1.19	1.19	1.19	1.09	1.06	1.12	1.08	0.96	1.07
Rand increase in wages	11.88	17.91	23.88	32.76	37.17	56.10	64.80	86.40	160.20

Source: NIDS Wave 1. Authors' calculations – design weights used

1/ Individuals who report a wage of "zero" are excluded from the analysis
3/ Average wage and transport cost of individuals in the given income decile
4/ Transport cost/Wage
5/ Bootstrapped standard errors used. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table 20b: Effect of Transport Costs on Wages by Income Decile, 2010 (Zero transport cost reporters excluded)

Percentile	10th	20th	30th	40th	Median	60th	70th	80th	90th
Wage 1/ 3/	1,831	1,657	3,062	3,321	3,702	4,770	6,198	10,292	10,421
Transport cost 3/	217	241	300	274	469	476	474	725	742
Proportion of income	0.12	0.14	0.10	0.08	0.13	0.10	0.08	0.07	0.07
Elasticity	0.349***	0.240***	0.242***	0.208***	0.202***	0.199***	0.197***	0.169***	0.154***
Marginal effect:									
6% increase in transport cost	12.76	13.77	18.00	16.47	28.14	28.56	28.47	44.68	44.52
% increase in wages	2.09	1.44	1.45	1.25	1.21	1.19	1.18	1.01	0.92
Rand increase in wages	20.94	21.60	29.04	37.44	42.42	59.70	70.92	91.26	138.60

Source: NIDS Wave 2. Authors' calculations – design weights used

1/ Individuals who report a wage of "zero" are excluded from the analysis
3/ Average wage and transport cost of individuals in the given income decile
4/ Transport cost/Wage
5/ Bootstrapped standard errors used. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Possible Sample Selection Bias: Non-randomness of Non-reporting

Out of a sample of 21,880 individuals, we can include only 2,567 individuals (or less) in our regression analysis in 2010.[23] In part, this reflects that only a subset of individuals fall into the category of being employed. In addition, of these, not all provide valid responses to the question on wages and transport cost. Out of 3,559 individuals responding to the wage question, only 3,038 gave a valid response (i.e. did not "refuse" or responded with "don't know."). And out of 2,738 individuals responding to the transport cost questions, only 2,567 were valid responses. We also note that 13 of the 3,038 who provide an answer to the wage question report a wage of "zero" while 522 of the 2,567 who reply to the transport costs question indicate a "zero" cost. It is plausible to label these responses as 'mis-reporting' as economic theory rules out zero wages while even those who walk all the way to work incur some form of transport costs (at the very least, the wear and tear of shoes).

Given that we have substantial non-reporting, the question becomes whether this non-reporting is systematic and thus may impact the interpretation of our regression results. We find more groups of individuals who respond with a "refuse" or "do not know" to the transport cost question. This becomes even more acute when we include potential transport cost mis-reporting, i.e. a reported transport cost of zero.[24] From the second column in Table A10 in the Appendix, it is evident that individuals residing in Mpumalanga, who are not part of a union and who reside in households low in per capita income, are less likely to report their transport costs whereas those in the wholesale and retail trade are more likely to report than individuals in the other industries. However, from the table below, we see that there does not appear to be substantial differences in the wages of the non-reporters compared to that of reporters.

Table 21: Comparison of Transport Costs and Wages of Non-reporters with Reporters, 2010 (rand per Month)

	Transport Costs		Wages	
	Non-reporters 1/	Reporters	Non-reporters 2/	Reporters
Median	400	380	5,000	4,000
Mean	556	493	5,272	5,954
Observations	179	2,380	132	2,377

Source: NIDS Wave 2. Authors' calculations using design weights

1/ Individuals who do not report a wage but report a transport cost - we note that there are some who report neither.
2/ Individuals who do not report a transport cost but report a wage

To ascertain that any potential bias is not significant, we use a Heckman selection estimator using the code of the enumerator and the time of interview as variables that explain the first stage (selection into not reporting either transport costs or reservation wages). Our choice of variables is based on the argument that "lazy" enumerators were more likely to skip questions and that

[23] Analysis conducted for 2008 as well and results are relatively similar – results available upon request from the authors

[24] Over and above those mentioned in the text, these include age, gender, education and location.

interviews conducted outside normal hours (08h00 to 17h00) were less likely to be conducted accurately.

The insignificant inverse mills ratio confirms that the observed, albeit slight, differences in Tables 21 and A10 are not cause for concern (see Table 22). Bearing in mind the sensitivity of the Heckman standard errors to the specification, the low degrees of freedom and the considerably reduced sample sizes; we opted to proceed with OLS estimation in the main text.

Table 22: Heckman Selection Estimation of the Effect of Transport Costs on Wages, 2010

Inverse mills ratio	-0.0522 (0.0470)
Observations	958
R-squared	0.532

Notes: Standard errors are clustered at the household level and sampling weights are included. Strata with single sampling unit are centered at the overall mean. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

VI. CONCLUSION

The legacy of apartheid's segregation policy is still visible with the geographic separation of large shares of the population and areas of economic activity. The extent to which this legacy affects labor market outcomes is of crucial importance in order to understand the South African labor market. In this paper, we have investigated the effect of transport costs on wages.

Our findings suggest that transport cost matter. We show that transport costs can devour a substantial part of a worker's wage at the lower end of the income distribution. Transport costs take up to 24 percent of an individual's actual wage at the lower end of the income distribution yet account for a mere 7 percent for the richest. This has implications for living standards and inequality – the Gini index for wage earnings rises markedly when one factors in transport costs. Workers try to recoup this cost which leads to wages being pushed up by increasing transport cost. However, the relationship between transport costs and reservation wages differs across the income distribution. Workers at the bottom of the income distribution seem to be able to recover a fraction of the transport costs they face, while workers at the top of the income distribution seem to be able to overcompensate for transport costs. This is likely to reflect differences in bargaining power, but possibly also the fact that at the top end, transport costs fade in comparison to wages.

An important question over the validity of our results is whether our data is truly representative of the South African labor market. We do not have an easy answer to this question, but worry that the transport costs, distances, and commute times in our sample are lower than what we would have expected from our own casual and anecdotal observations.

This question notwithstanding, there are some tentative policy implications that one can draw from our results. Firstly, transport costs matter for wages, but, in terms of magnitude, possibly

more at the upper end of the wage distribution where unemployment—largely a problem for the unskilled—is less of a problem. As such, spatial distances through their impact on wages, may not be a major factor explaining South Africa's high unemployment—assuming that our sample is also representative of South Africa's unemployed. Even in this case, the impact of transport cost on search activity may be a factor, one that we have not studied in this paper. Second, at the lower end of the income distribution, transport costs devour a large share of the take-home wage and thus contribute to poverty and income inequality. Finding ways to reduce transport costs—through better public transport or by moving work places closer to where people live—should, therefore, help tackle South Africa's triple challenges.

REFERENCES

Arora, V., and Ricci, L. (2005). Unemployment and the Labour Market. In *Post-apartheid South Africa: The First Ten Years.* International Monetary Fund.

Banerjee, A., Galiani, S., Levinsohn, J., McLaren, Z., and Woolard, I. (2008). Why has Unemployment Risen in the New South Africa? *The Economics of Transition, 16*(4), 715–740.

Bloemen, H. G., and Stancanelli, E. G. (2001). Individual Wealth, Reservation Wages, and Transitions into Employment. *Journal of Labor Economics, 19*(2), 400-439.

Carolina, M. G., and Pau, L. (2008). *The Reservation Wage of Unemployed Job Seekers in Curaçao.* Central Bank of the Netherlands Antilles.

Church, A., Frost, M., and Sullivan, K. (2000). Transport and Social Exclusion in London. *Transport Policy, 7*(3), 195-205.

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

De Lange, K. (2013). *The Potential Effects of Transport Costs on Unemployment in South Africa.*

Dimitrov, L. (2010). The Effects of Social Exclusion and Transport in South Africa. *29th Southern African Transport Conference*, (pp. 723-732). Pretoria.

The Economist. (2014, October 25). *The Geography of Joblessness - The Difficulty People have in Getting to Jobs Makes Unemployment Unnecessarily High*. Retrieved from http://www.economist.com/news/finance-and-economics/21627628-difficulty-people-have-getting-jobs-makes-unemployment-unnecessarily

Fourie, F. C. (2011). *The South African Unemployment Debate: Three Worlds, Three Discourses?* University of Cape Town, Southern Africa Labour and Development Research Unit.

Franklin, S. (2014). Job Search, Transport Costs and Youth Unemployment: Evidence from Urban Ethiopia.

Heintz, J., and Posel, D. (2008). Revisiting Informal Unemployment and Segmentation in the South African Labour Market. *South Africa Journal of Economics, 76*(1), 26-44.

Hinks, T. (2008). Poverty, Networks and Location Determinants of Job Search in South Africa. *Journal of International Development, 20*, 117-131.

Holzer, H. J., Ihlanfeldt, K. R., and Sjoquist, D. (1994). Work, Search, and Travel among White and Black Youth. *Journal of Urban Economics, 35*, 320-345.

Kingdon, G., and Knight, J. (2006). The Measurement of Unemployment when Unemployment is High. *Labour Economics, 13*, 291-315.

Koenker, R., and Bassett, G. (1978). Regression Quantiles. *Econometrica, 46*(1), 33-50.

Krueger, A. B., and Mueller, A. (2011). Job Search, Emotional Well-Being and Job Finding in a Period of Mass Unemployment: Evidence from High-Frequency Longitudinal Data. *Brookings Papers on Economic Activity, 42*(1), 1-81.

Krueger, A. B., and Mueller, A. (2014). *A Contribution to the Empirics of Reservation Wages.* Cambridge: National Bureau of Economic Research.

Laird, J. (2006). *Commuting Costs and their Impact on Wage Rates.* University of Leeds, Institute of Transport Studies.

Lammers, M. (2009). The Effects of Savings on Reservation Wages and Search Effort.

Maani, S. A., and Studenmund, A. H. (1986). The Critical Wage, Unemployment Duration, and Wage Expectations: The Case of Chile. *Industrial and Labor Relations Review, 39*(2), 264-276.

Mahajan, S. (2011, March 5). *After the Long Walk to Freedom, an Affordable Bus Ride to Work*. Retrieved from World Bank: http://blogs.worldbank.org/africacan/after-the-long-walk-to-freedom-an-affordable-bus-ride-to-work.

Mortensen, D. (1986). Job Search and Labor Market Analysis. In O. Ashenfelter, and R. Layard.

Nattrass, N. (2000). The Debate about Unemployment in the 1990s. *Studies in Economics and Econometrics, 24*(3), 73-89.

Nattrass, N., and Walker, R. (2005). Unemployment and Reservation Wages in Working-Class Cape Town. *South Africa Journal of Economics, 73*(3), 498-509.

Patacchini, E., and Zenou, Y. (2005). Spatial Mismatch, Transport Mode and Search Decisions in England. *Journal of Urban Economics, 58*(1), 62-90.

Phillips, D. (2011). *Getting to Work: Experimental Evidence on Job Search and Transportation Costs.* National Poverty Center.

Prasad, E. S. (2003). *What Determines the Reservation Wages of Unemployed Workers? New Evidence from German Micro Data.* Institute for the Study of Labor (IZA).

Rankin, N., and Roberts, G. (2010). *Youth Unemployment, Firm Size, and Reservation Wages in South Africa.* Munich Personal RePEc Archive.

Rodrik, D. (2006). *Understanding South Africa's Economic Puzzles.* Centre for International Development at Havard University.

Rouwendal, J., and Rietveld, P. (1994). Changes in Commuting Distance of Dutch Households. *Urban Studeis, 31,* 1545-1557.

Rupert, P., Stancanelli, E., and Wasmer, E. (2009). *Commuting, Wages and Bargaining Power.* University Cergy.

Smith, T. E., and Zenou, Y. (2003). Spatial Mismatch, Search Effort, and Urban Spatial Structure. *Journal of Urban Economics, 54(1),* 129-156.

Statistics South Africa (StatsSA). (2014). *National Household Transport Survery.* Department of Transport.

Van den Berg, G. J., and Gorter, C. (1997). Job Search and Commuting Time. *Journal of Business and Economics Statistics, 15*(2), 269-281.

Van Ommeren, J. (1996). *Commuting and Relocation of Jobs and Residences; A Search Perspective.*

Van Ommeren, J., and Fosgerau, M. (2009). Workers' Marginal Costs of Commuting. *Journal of Urban Economics, 65(1),* 38-47.

Verick, S. (2011). *Giving Up Job Search During a Recession: The Impact of the Global Financial Crisis on the South African Labour Market.* Institute for the Study of Labor (IZA).

Walker, R. (2003). *Reservation Wages - Measurement and Determinants: Evidence from the Khayelitsha/ Mitchell's Plain Survey.* University of Cape Town, Centre for Social Science Research.

White, M. J. (1986). Sex Differences in Urban Commuting Patterns. *The American Economic Review, 76(2),* 368-372.

APPENDIX

Table A1: Effect of Transport Costs on Wages, 2008

	OLS		IV		HECK + IV	
Log of wages	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)
Log of transport costs	0.0459***	0.184***	-0.140*	0.675*	-0.245	0.453*
	(0.0061)	(0.0303)	(0.0735)	(0.346)	(0.203)	(0.245)
Age	0.00693***	0.00748***	0.0139***	0.00879***	0.0185*	0.00378
	(0.00167)	(0.00205)	(0.00339)	(0.00258)	(0.00973)	(0.00370)
Male	0.205***	0.0865**	0.228***	0.0748		
	(0.0391)	(0.043)	(0.0388)	(0.0552)		
Married	0.144***	0.123***	0.0730	0.0343	-0.0947	-0.124
	(0.038)	(0.0454)	(0.0468)	(0.0770)	(0.108)	(0.0965)
Race						
African						
Coloured	0.238***	0.227**	0.187***	0.137	-0.0798	-0.00403
	(0.0806)	(0.0997)	(0.0670)	(0.0969)	(0.180)	(0.139)
Asian/Indian	0.0872	0.197**	0.404**	-0.0300	0.942	-0.0135
	(0.104)	(0.0985)	(0.164)	(0.203)	(0.633)	(0.257)
White	0.458***	0.429***	0.614***	0.204	0.917***	0.282
	(0.0677)	(0.0719)	(0.0851)	(0.171)	(0.301)	(0.181)
Highest level of education attained						
No schooling						
Primary	-0.0172	-0.118	0.0292	-0.141	0.0955	0.00260
	(0.065)	(0.111)	(0.0745)	(0.121)	(0.154)	(0.146)
Incomplete secondary	0.125*	0.0224	0.265***	0.0910	0.186	0.164
	(0.0718)	(0.108)	(0.0858)	(0.111)	(0.150)	(0.137)
Matric	0.470***	0.345***	0.679***	0.267**	0.332	0.335
	(0.0827)	(0.114)	(0.109)	(0.130)	(0.236)	(0.222)
Certificate	0.0973	0.0896	0.134*	0.00795	-0.111	-0.103
	(0.0631)	(0.0776)	(0.0712)	(0.0905)	(0.160)	(0.132)
Diploma	0.308***	0.261***	0.365***	0.188**	-0.159	0.0910
	(0.0687)	(0.0832)	(0.0664)	(0.0845)	(0.259)	(0.186)
Degree	0.641***	0.593***	0.859***	0.368**	0.394*	0.365
	(0.0791)	(0.0911)	(0.121)	(0.179)	(0.230)	(0.243)
Household composition						
Household income per capita	0.0523***	0.0572***	0.00339	0.0682***	-0.103	0.0311
	(0.00766)	(0.00932)	(0.0209)	(0.0145)	(0.0824)	(0.0287)
Main breadwinner						
Occupation category						
Elementary occupations						
Armed forces						
Managers	-0.0846	0.0219	-0.103	0.107	-0.0352	0.0657

	(0.0973)	(0.0955)	(0.0955)	(0.124)	(0.208)	(0.160)
Professionals	-0.131	-0.051	-0.191*	0.0988	-0.155	0.0170
	(0.128)	(0.146)	(0.105)	(0.157)	(0.219)	(0.183)
Technicians and associate professionals	-0.237**	-0.192	-0.121	-0.168	0.101	-0.235
	(0.0999)	(0.117)	(0.101)	(0.106)	(0.272)	(0.152)
Clerical support workers	-0.409***	-0.278**	-0.593***	-0.121	-1.081**	-0.0865
	(0.0932)	(0.108)	(0.109)	(0.171)	(0.496)	(0.219)
Service and sales workers	-0.492***	-0.00361	-0.856***	0.249	-1.098**	0.310
	(0.131)	(0.29)	(0.179)	(0.313)	(0.514)	(0.414)
Skilled agricultural, forestry and fishery workers	-0.352***	-0.333***	-0.519***	-0.204	-0.839***	-0.597**
	(0.0921)	(0.103)	(0.106)	(0.159)	(0.260)	(0.236)
Craft and related trades workers	-0.303***	-0.259**	-0.516***	-0.132	-0.688**	-0.545***
	(0.098)	(0.111)	(0.123)	(0.172)	(0.269)	(0.209)
Plant and machine operators, and assemblers	-0.543***	-0.523***	-0.806***	-0.302	-1.112***	-0.587***
	(0.0898)	(0.108)	(0.130)	(0.228)	(0.347)	(0.208)
Industry						
Community, social and personal services						
Private households	0.0669	0.0445	-0.0281	0.0707	-0.306	-0.0396
	(0.0895)	(0.172)	(0.0973)	(0.188)	(0.211)	(0.181)
Agriculture, hunting, forestry and fishing	0.710***	0.836***	0.708***	0.564**	1.182***	0.746**
	(0.118)	(0.152)	(0.108)	(0.238)	(0.395)	(0.309)
Mining and quarrying	0.244***	0.327***	0.439***	0.167	0.223	0.188
	(0.0688)	(0.0951)	(0.115)	(0.150)	(0.149)	(0.142)
Manufacturing	0.463***	0.352*	0.305	0.480*	-0.249	0.623
	(0.172)	(0.186)	(0.201)	(0.270)	(1.123)	(0.878)
Electricity, gas and water supply	0.166	0.0768	0.103	-0.109	-0.329	0.513
	(0.106)	(0.148)	(0.0988)	(0.167)	(0.568)	(0.334)
Construction	0.208***	0.14	0.239***	-0.0207	0.237	-0.0347
	(0.0711)	(0.0951)	(0.0813)	(0.142)	(0.166)	(0.147)
Wholesale and retail trade; repair; hotels	0.465***	0.432***	0.346***	0.141	0.316	0.313
	(0.0845)	(0.126)	(0.107)	(0.237)	(0.253)	(0.208)
Transport, storage and communication	0.322***	0.364***	0.279***	0.305***	0.0755	0.177
	(0.0723)	(0.099)	(0.0796)	(0.110)	(0.155)	(0.156)
Financial intermediation, insurance, real estate	0.339***	0.270***	0.175*	0.204*	-0.0107	0.182
	(0.0701)	(0.0906)	(0.0921)	(0.108)	(0.163)	(0.112)
Union membership	0.275***	0.237***	0.387***	0.223***	0.505***	0.241***
	(0.0388)	(0.046)	(0.0601)	(0.0468)	(0.145)	(0.0666)
Regional location of household						
Urban formal						
Rural formal	-0.182***	-0.092	-0.0257	-0.0882	0.444	0.340
	(0.0637)	(0.087)	(0.0881)	(0.125)	(0.433)	(0.285)
Tribal authority areas	0.073	0.0808	0.392***	0.0422	0.735	0.475**
	(0.0551)	(0.0815)	(0.133)	(0.123)	(0.478)	(0.221)
Urban informal	-0.277***	-0.221**	0.205	-0.193	0.621	0.219

	(0.0786)	(0.0963)	(0.195)	(0.125)	(0.602)	(0.278)
Province						
KwaZulu-Natal						
Western Cape	-0.164*	-0.205**	-0.151*	-0.226**	0.0531	-0.384***
	(0.0838)	(0.104)	(0.0775)	(0.0989)	(0.200)	(0.140)
Eastern Cape	-0.0458	0.0369	-0.298**	0.0226	-0.213	-0.00982
	(0.0906)	(0.131)	(0.151)	(0.194)	(0.300)	(0.254)
Northern Cape	-0.162	-0.162	-0.207**	-0.0301	-0.458**	-0.209
	(0.101)	(0.125)	(0.0867)	(0.127)	(0.225)	(0.134)
Free State	-0.0256	-0.0782	-0.0411	-0.129	0.00183	-0.169
	(0.0786)	(0.0879)	(0.0728)	(0.0984)	(0.132)	(0.112)
North West	0.0587	-0.125	-0.00602	-0.254*	-0.0837	-0.523***
	(0.0834)	(0.0994)	(0.0889)	(0.132)	(0.233)	(0.157)
Gauteng	0.104	0.0297	0.184***	-0.0588	0.401**	-0.0397
	(0.078)	(0.0881)	(0.0688)	(0.0933)	(0.195)	(0.109)
Mpumalanga	0.0675	0.00609	0.0657	-0.0687	0.149	-0.115
	(0.0812)	(0.0823)	(0.0769)	(0.110)	(0.178)	(0.129)
Limpopo	0.0549	0.0715	0.0263	0.0936	-0.0745	-0.133
	(0.084)	(0.11)	(0.0846)	(0.130)	(0.170)	(0.178)
Inverse mills ratio					-0.858	-0.153
					(0.533)	(0.343)
Constant	6.429***	5.793***	6.784***	3.110	8.513***	4.364***
	(0.163)	(0.254)	(0.214)	(1.926)	(1.082)	(1.646)
Observations	2,481	1,233	2,328	1,149	997	548
R-squared	0.671	0.682	0.422	0.536	0.255	0.710

Notes: Standard errors are clustered at the household level and sampling weights are included. Strata with single sampling unit are centered at the overall mean. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table A2: Effect of Transport Costs on Wages, 2010

	OLS		IV		HECK + IV	
Log of wages	(1a)	(1b)	(2a)	(2b)	(3a)	(3b)
Log of transport costs	0.0263**	0.181***	0.411*	1.523	0.248	0.836
	(0.0114)	(0.0344)	(0.231)	(1.924)	(0.261)	(1.132)
Age	0.0100***	0.00719***	-0.00207	-0.0109	0.0209***	0.0172***
	(0.00258)	(0.00269)	(0.00781)	(0.0269)	(0.00642)	(0.00592)
Male	0.268***	0.233***	0.443***	0.0273		
	(0.0560)	(0.0573)	(0.115)	(0.321)		
Married	0.133***	0.0976**	0.0633	0.0743	-0.0116	0.0204
	(0.0459)	(0.0488)	(0.0679)	(0.0749)	(0.0613)	(0.102)
Race						
African						
Coloured	0.352***	0.442***	0.300***	0.394***	0.0942	0.0957
	(0.0927)	(0.0945)	(0.105)	(0.148)	(0.190)	(0.334)
Asian/Indian	0.567***	0.491***	0.657***	-0.0620	0.542*	0.539
	(0.121)	(0.127)	(0.178)	(0.834)	(0.289)	(0.427)
White	0.611***	0.657***	0.606***	0.486*	0.953***	0.785***
	(0.0835)	(0.0768)	(0.0878)	(0.260)	(0.148)	(0.254)
Age						
Highest level of education attained						
No schooling						
Primary	-0.0207	0.0512	-0.0866	0.341	-0.229	-0.232
	(0.159)	(0.215)	(0.160)	(0.464)	(0.173)	(0.194)
Incomplete secondary	0.156	0.190	-0.116	0.0469	-0.366	-0.357
	(0.151)	(0.199)	(0.217)	(0.291)	(0.253)	(0.272)
Matric	0.289*	0.288	-0.103	0.0392	-0.693*	-0.586*
	(0.151)	(0.196)	(0.277)	(0.429)	(0.368)	(0.308)
Certificate	0.181**	0.171**	-0.0235	0.117	-0.370	-0.211
	(0.0753)	(0.0758)	(0.149)	(0.132)	(0.307)	(0.193)
Diploma	0.386***	0.298***	0.468***	0.218	-0.606	-0.516
	(0.0922)	(0.0898)	(0.0953)	(0.179)	(0.490)	(0.423)
Degree	0.687***	0.646***	0.362	0.111	-0.780	-0.688
	(0.0930)	(0.0961)	(0.229)	(0.799)	(0.768)	(0.734)
Household composition						
Household income per capita	0.0576***	0.0563***	0.0858***	0.0922	0.00557	0.00873
	(0.0111)	(0.0120)	(0.0219)	(0.0570)	(0.0257)	(0.0276)
Main breadwinner	-0.0747	-0.0603	0.0203	0.0655	-0.236	-0.178
	(0.0536)	(0.0547)	(0.0833)	(0.195)	(0.152)	(0.136)
Occupation category						
Elementary occupations						
Armed forces	0.877***	0.850***	0.300	0.152		
	(0.150)	(0.159)	(0.566)	(1.104)		
Managers	0.862***	0.780***	0.552**	0.149	0.857***	0.805***

	(0.131)	(0.122)	(0.226)	(0.927)	(0.189)	(0.279)
Professionals	0.531***	0.441***	0.401***	-0.0719	0.239	0.153
	(0.106)	(0.0991)	(0.136)	(0.751)	(0.218)	(0.395)
Technicians and associate professionals	0.576***	0.523***	0.249	-0.150	0.293	0.294
	(0.127)	(0.119)	(0.228)	(0.973)	(0.233)	(0.256)
Clerical support workers	0.324***	0.321***	0.237*	-0.0492	0.138	0.145
	(0.111)	(0.105)	(0.129)	(0.575)	(0.194)	(0.233)
Service and sales workers	0.0922	0.0981	0.0588	-0.0145	0.0332	0.0858
	(0.0735)	(0.0757)	(0.0973)	(0.222)	(0.114)	(0.122)
Skilled agricultural, forestry and fishery workers	-1.565**	-1.593**	-1.753***	-1.166	-0.228	0.0876
	(0.610)	(0.636)	(0.330)	(0.995)	(0.673)	(0.914)
Craft and related trades workers	0.0755	0.00463	-0.0352	-0.229	0.0808	0.109
	(0.0881)	(0.0922)	(0.120)	(0.353)	(0.204)	(0.234)
Plant and machine operators, and assemblers	0.269***	0.252***	0.355***	0.370*	0.571*	0.364
	(0.0803)	(0.0859)	(0.127)	(0.222)	(0.308)	(0.304)
Industry						
Community, social and personal services						
Private households	-0.321***	-0.290***	-0.580***	-0.261	-0.535***	-0.397***
	(0.103)	(0.0992)	(0.199)	(0.168)	(0.158)	(0.141)
Agriculture, hunting, forestry and fishing	-0.325***	-0.171	0.684	0.274	0.293	-0.467
	(0.112)	(0.182)	(0.636)	(0.691)	(0.903)	(0.326)
Mining and quarrying	0.327***	0.334***	0.132	0.0683	0.601*	0.565*
	(0.105)	(0.102)	(0.191)	(0.467)	(0.338)	(0.297)
Manufacturing	-0.0740	-0.115	-0.246	-0.178	-0.641***	-0.641***
	(0.118)	(0.103)	(0.155)	(0.168)	(0.151)	(0.172)
Electricity, gas and water supply	0.249	0.418**	0.354	0.648	0.884**	0.787*
	(0.191)	(0.165)	(0.347)	(0.560)	(0.400)	(0.465)
Construction	-0.0268	0.00642	0.0978	0.179	-0.359	-0.596
	(0.110)	(0.117)	(0.147)	(0.289)	(0.363)	(0.713)
Wholesale and retail trade; repair; hotels	-0.0868	-0.0918	-0.0763	-0.206	-0.171	-0.160
	(0.0839)	(0.0743)	(0.0857)	(0.214)	(0.118)	(0.182)
Transport, storage and communication	0.0355	0.0396	-0.107	-0.0799	0.889***	0.0937
	(0.114)	(0.125)	(0.167)	(0.236)	(0.331)	(0.775)
Financial intermediation, insurance, real estate	0.0697	0.0557	-0.0706	-0.0254	0.00935	-0.144
	(0.0761)	(0.0739)	(0.130)	(0.169)	(0.124)	(0.358)
Union membership	0.165***	0.148***	0.0446	0.0194	0.160	0.0709
	(0.0509)	(0.0542)	(0.0890)	(0.182)	(0.0988)	(0.234)
Regional location of household						
Urban formal						
Rural formal	-0.166*	-0.164	0.0504	-0.315	-0.0902	-0.490
	(0.0944)	(0.118)	(0.174)	(0.244)	(0.219)	(0.345)
Tribal authority areas	-0.155**	-0.130*	-0.114	-0.170	0.649**	0.418*
	(0.0743)	(0.0735)	(0.103)	(0.135)	(0.324)	(0.226)
Urban informal	-0.0549	0.0329	0.000574	0.153	0.0709	0.187

	(0.0848)	(0.0766)	(0.0920)	(0.193)	(0.101)	(0.172)
Province						
KwaZulu-Natal						
Western Cape	0.141	0.130	0.460**	0.195	0.276	0.288
	(0.121)	(0.125)	(0.223)	(0.189)	(0.185)	(0.202)
Eastern Cape	-0.144	-0.146	0.00746	0.0417	0.0618	0.296
	(0.106)	(0.0978)	(0.142)	(0.313)	(0.138)	(0.389)
Northern Cape	-0.0363	0.0489	0.836	0.205	0.903	0.696
	(0.117)	(0.113)	(0.553)	(0.389)	(0.718)	(0.559)
Free State	-0.0398	0.0503	0.180	0.233	0.221	0.327*
	(0.102)	(0.0948)	(0.160)	(0.288)	(0.148)	(0.177)
North West	0.231**	0.218**	0.142	-0.0875	-0.195	-0.0129
	(0.107)	(0.102)	(0.153)	(0.496)	(0.249)	(0.190)
Gauteng	0.112	0.101	-0.0167	-0.224	0.177	0.201
	(0.0853)	(0.0795)	(0.120)	(0.486)	(0.125)	(0.165)
Mpumalanga	0.141	0.0760	0.225*	-0.177	-0.311	-0.240
	(0.102)	(0.105)	(0.123)	(0.400)	(0.303)	(0.291)
Limpopo	0.0183	0.0183	0.00411	0.144	-0.178	0.229
	(0.0911)	(0.102)	(0.133)	(0.262)	(0.209)	(0.373)
Inverse mills ratio					-1.720**	-1.306**
					(0.780)	(0.511)
Constant	6.431***	5.673***	4.922***	-0.951	7.766***	3.446
	(0.223)	(0.302)	(0.943)	(9.452)	(0.781)	(6.588)
Observations	1,894	1,496	1,799	1,417	826	689
R-squared	0.578	0.596	0.010	-0.233	0.534	0.485

Notes: Standard errors are clustered at the household level and sampling weights are included. Strata with single sampling unit are centered at the overall mean. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table A3: Effect of Transport Cost on Wages by Income Decile, 2008

	Percentile								
Log of wages	10th	20th	30th	40th	Median	60th	70th	80th	90th
Log of transport costs	0.349***	0.240***	0.242***	0.208***	0.202***	0.199***	0.197***	0.169***	0.154***
	(0.0645)	(0.0460)	(0.0358)	(0.0328)	(0.0307)	(0.0273)	(0.0253)	(0.0255)	(0.0338)
Age	0.00738	0.00874**	0.00829***	0.00782***	0.00907***	0.00761***	0.0104***	0.0110***	0.0153***
	(0.00479)	(0.00340)	(0.00261)	(0.00237)	(0.00211)	(0.00213)	(0.00223)	(0.00270)	(0.00285)
Male	0.0989	0.218***	0.183***	0.221***	0.200***	0.223***	0.202***	0.175***	0.165**
	(0.100)	(0.0679)	(0.0572)	(0.0475)	(0.0431)	(0.0408)	(0.0432)	(0.0502)	(0.0678)
Married	0.0548	0.0795	0.0377	0.0367	0.0413	0.0539	0.0870**	0.126***	0.108*
	(0.0791)	(0.0584)	(0.0468)	(0.0430)	(0.0403)	(0.0406)	(0.0398)	(0.0458)	(0.0585)
Race									
African									
Coloured	0.356*	0.220*	0.169*	0.118	0.147**	0.186***	0.229***	0.262***	0.220**
	(0.198)	(0.120)	(0.102)	(0.0845)	(0.0729)	(0.0685)	(0.0699)	(0.0735)	(0.0946)
Asian/Indian	0.510*	0.636***	0.364**	0.305**	0.231	0.238	0.351*	0.359**	0.332**
	(0.307)	(0.218)	(0.164)	(0.142)	(0.156)	(0.190)	(0.182)	(0.158)	(0.153)
White	0.595***	0.582***	0.598***	0.536***	0.599***	0.643***	0.651***	0.564***	0.410***
	(0.172)	(0.128)	(0.105)	(0.1000)	(0.0871)	(0.0803)	(0.0808)	(0.0825)	(0.0979)
Highest level of education attained									
No schooling									
Primary	0.108	0.127	0.0685	0.0118	0.0538	0.0354	0.0493	0.0809	-0.108
	(0.221)	(0.154)	(0.136)	(0.131)	(0.105)	(0.0944)	(0.104)	(0.125)	(0.142)
Incomplete secondary	0.170	0.107	0.117	0.0662	0.0560	0.0670	0.0713	0.117	0.00429
	(0.199)	(0.150)	(0.134)	(0.121)	(0.0949)	(0.0852)	(0.0968)	(0.118)	(0.140)
Matric	0.343	0.385**	0.356**	0.239*	0.242**	0.267***	0.277***	0.338***	0.270*
	(0.220)	(0.170)	(0.148)	(0.128)	(0.0997)	(0.0932)	(0.104)	(0.125)	(0.140)
Certificate	0.0412	0.0876	0.112	0.128*	0.108*	0.0900	0.0922	0.0975	0.102
	(0.156)	(0.103)	(0.0820)	(0.0744)	(0.0651)	(0.0609)	(0.0631)	(0.0775)	(0.113)
Diploma	0.263**	0.266***	0.334***	0.305***	0.287***	0.326***	0.303***	0.295***	0.248**
	(0.132)	(0.103)	(0.0914)	(0.0774)	(0.0726)	(0.0693)	(0.0688)	(0.0734)	(0.0975)

Degree	0.278	0.477***	0.595***	0.573***	0.514***	0.468***	0.433***	0.540***	0.567***
	(0.244)	(0.159)	(0.124)	(0.114)	(0.0939)	(0.0895)	(0.102)	(0.115)	(0.128)
Household composition									
Household income per capita	0.0866***	0.0416***	0.0276**	0.0272**	0.0361***	0.0326***	0.0346***	0.0358***	0.0374***
	(0.0195)	(0.0145)	(0.0119)	(0.0120)	(0.0106)	(0.00944)	(0.00948)	(0.0105)	(0.0128)
Main breadwinner	-0.0532	-0.0598	-0.0142	-0.0104	-0.0204	-0.0216	-0.0595	-0.0546	-0.114**
	(0.0990)	(0.0714)	(0.0520)	(0.0439)	(0.0407)	(0.0402)	(0.0425)	(0.0513)	(0.0573)
Occupation category									
Elementary occupations									
Armed forces	0.769**	0.794**	1.166***	1.033***	1.059***	0.940***	0.837***	0.869***	0.835***
	(0.387)	(0.398)	(0.383)	(0.320)	(0.271)	(0.246)	(0.213)	(0.203)	(0.221)
Managers	0.772***	0.676***	0.751***	0.838***	0.848***	0.768***	0.927***	0.959***	0.969***
	(0.216)	(0.155)	(0.143)	(0.122)	(0.110)	(0.131)	(0.145)	(0.147)	(0.316)
Professionals	0.421**	0.494***	0.432***	0.540***	0.611***	0.511***	0.509***	0.517***	0.533***
	(0.186)	(0.123)	(0.107)	(0.102)	(0.0920)	(0.0841)	(0.0855)	(0.0946)	(0.102)
Technicians and associate professionals	0.168	0.308*	0.331**	0.608***	0.666***	0.520***	0.606***	0.587***	0.673***
	(0.288)	(0.180)	(0.166)	(0.150)	(0.110)	(0.0971)	(0.0936)	(0.114)	(0.177)
Clerical support workers	0.172	0.458***	0.343***	0.512***	0.531***	0.426***	0.398***	0.380***	0.460***
	(0.246)	(0.149)	(0.104)	(0.0926)	(0.0806)	(0.0752)	(0.0758)	(0.0976)	(0.118)
Service and sales workers	0.0900	0.0808	0.109	0.178***	0.191***	0.161***	0.193***	0.207***	0.206***
	(0.128)	(0.0882)	(0.0763)	(0.0647)	(0.0596)	(0.0625)	(0.0589)	(0.0627)	(0.0788)
Skilled agricultural, forestry and fishery workers	-5.189**	-1.009	-0.936	-0.0439	-0.0487	-0.136	-0.216	-0.323	-0.258
	(2.470)	(2.159)	(1.551)	(0.963)	(0.624)	(0.353)	(0.273)	(0.294)	(0.258)
Craft and related trades workers	-0.138	-0.00926	0.104	0.118	0.147*	0.138*	0.121	0.150*	0.136
	(0.196)	(0.160)	(0.109)	(0.0848)	(0.0792)	(0.0773)	(0.0771)	(0.0878)	(0.109)
Plant and machine operators, and assemblers	0.201	0.183	0.255***	0.266***	0.258***	0.224***	0.336***	0.333***	0.299***
	(0.227)	(0.151)	(0.0929)	(0.0875)	(0.0770)	(0.0794)	(0.0822)	(0.0777)	(0.0852)
Industry									
Community, social and personal services									
Private households	-0.358**	-0.264**	-0.300***	-0.380***	-0.354***	-0.447***	-0.433***	-0.489***	-0.544***
	(0.166)	(0.119)	(0.0966)	(0.0894)	(0.0846)	(0.0798)	(0.0815)	(0.0830)	(0.0978)

Agriculture, hunting, forestry and fishing	-0.471	0.0682	-0.0527	-0.0592	-0.167	-0.225**	-0.198	-0.247	0.0511
	(0.372)	(0.263)	(0.169)	(0.137)	(0.118)	(0.113)	(0.123)	(0.159)	(0.235)
Mining and quarrying	0.310	0.380**	0.419***	0.407***	0.389***	0.351***	0.327***	0.309***	0.226**
	(0.240)	(0.164)	(0.123)	(0.104)	(0.0892)	(0.0931)	(0.0985)	(0.102)	(0.108)
Manufacturing	-0.411**	-0.177	-0.200*	-0.168	-0.124	-0.166**	-0.130*	-0.102	-0.0891
	(0.202)	(0.126)	(0.105)	(0.104)	(0.0837)	(0.0789)	(0.0767)	(0.0818)	(0.0944)
Electricity, gas and water supply	0.595*	0.628**	0.632**	0.443*	0.380*	0.315	0.192	0.239	1.270*
	(0.329)	(0.288)	(0.271)	(0.236)	(0.212)	(0.256)	(0.376)	(0.534)	(0.676)
Construction	-0.0841	-0.0129	-0.0853	-0.118	-0.0998	-0.0902	0.0869	0.115	0.185
	(0.283)	(0.169)	(0.135)	(0.119)	(0.129)	(0.149)	(0.138)	(0.140)	(0.188)
Wholesale and retail trade; repair; hotels	-0.264*	-0.153	-0.120	-0.155**	-0.153**	-0.241***	-0.203***	-0.228***	-0.239***
	(0.143)	(0.0982)	(0.0814)	(0.0735)	(0.0651)	(0.0638)	(0.0642)	(0.0700)	(0.0796)
Transport, storage and communication	-0.209	0.145	0.131	0.0346	0.00651	-0.0248	-0.0984	-0.0383	0.118
	(0.372)	(0.178)	(0.107)	(0.0961)	(0.0981)	(0.101)	(0.115)	(0.144)	(0.142)
Financial intermediation, insurance, real estate	0.142	0.0510	0.0295	-0.0356	-0.0244	-0.0831	-0.0929	-0.131*	-0.200*
	(0.129)	(0.0950)	(0.0847)	(0.0780)	(0.0718)	(0.0642)	(0.0655)	(0.0731)	(0.102)
Union membership	0.157	0.241***	0.231***	0.177***	0.197***	0.191***	0.190***	0.160***	0.135**
	(0.120)	(0.0788)	(0.0589)	(0.0496)	(0.0464)	(0.0473)	(0.0454)	(0.0451)	(0.0554)
Regional location of household									
Urban formal									
Rural formal	-0.332	-0.251*	-0.137	-0.159*	-0.136*	-0.110	-0.118*	-0.125*	-0.187**
	(0.217)	(0.152)	(0.115)	(0.0846)	(0.0734)	(0.0695)	(0.0654)	(0.0694)	(0.0885)
Tribal authority areas	-0.202	-0.157	-0.189**	-0.193***	-0.156***	-0.166***	-0.170***	-0.149**	-0.124*
	(0.140)	(0.101)	(0.0768)	(0.0722)	(0.0599)	(0.0560)	(0.0556)	(0.0620)	(0.0717)
Urban informal	0.0639	0.0626	-0.0588	-0.0469	-0.0687	-0.0695	-0.0754	-0.109	-0.0611
	(0.117)	(0.0848)	(0.0717)	(0.0627)	(0.0583)	(0.0578)	(0.0577)	(0.0704)	(0.0871)
Province									
KwaZulu-Natal									
Western Cape	0.256	0.365**	0.288**	0.288***	0.225***	0.174**	0.0718	-0.0143	-0.0780
	(0.267)	(0.161)	(0.130)	(0.110)	(0.0802)	(0.0713)	(0.0736)	(0.0766)	(0.0853)
Eastern Cape	0.105	-0.0317	-0.0756	-0.126	-0.182**	-0.127	-0.0869	-0.136*	-0.152

	(0.152)	(0.130)	(0.111)	(0.0998)	(0.0853)	(0.0787)	(0.0780)	(0.0796)	(0.107)
Northern Cape	0.386*	0.304*	0.201	0.152	0.0573	0.0617	-0.00319	-0.00773	0.101
	(0.226)	(0.177)	(0.149)	(0.128)	(0.103)	(0.0890)	(0.0918)	(0.124)	(0.191)
Free State	0.194	0.162	0.0419	-0.00181	-0.0873	-0.0508	-0.0899	-0.0452	-0.0352
	(0.192)	(0.135)	(0.109)	(0.103)	(0.0794)	(0.0761)	(0.0919)	(0.0961)	(0.103)
North West	0.529**	0.413***	0.258***	0.158*	0.0998	0.101	0.123	0.119	0.139
	(0.211)	(0.122)	(0.0932)	(0.0929)	(0.0807)	(0.0809)	(0.0808)	(0.0831)	(0.101)
Gauteng	0.295*	0.261**	0.191**	0.141*	0.0694	0.0565	0.0434	0.0462	0.0634
	(0.152)	(0.110)	(0.0823)	(0.0798)	(0.0617)	(0.0577)	(0.0622)	(0.0679)	(0.0890)
Mpumalanga	0.159	0.256**	0.173*	0.123	0.0137	-0.0140	0.0211	-0.0248	0.0930
	(0.161)	(0.127)	(0.0946)	(0.0891)	(0.0665)	(0.0678)	(0.0755)	(0.0866)	(0.114)
Limpopo	0.0885	0.0787	0.174	0.0688	-0.0199	-0.00351	0.0939	0.143	0.0592
	(0.217)	(0.185)	(0.109)	(0.0927)	(0.0836)	(0.0896)	(0.0879)	(0.0962)	(0.0896)
Constant	3.910***	4.770***	5.120***	5.604***	5.730***	5.984***	5.999***	6.238***	6.516***
	(0.480)	(0.326)	(0.276)	(0.242)	(0.213)	(0.202)	(0.211)	(0.238)	(0.277)
Observations	1,510	1,510	1,510	1,510	1,510	1,510	1,510	1,510	1,510

Notes: Standard errors are clustered at the household level and sampling weights are included. Strata with single sampling unit are centered at the overall mean.
Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table A4: Effect of Transport Cost on Wages by Income Decile, 2010

	Percentile								
Log of wages	10th	20th	30th	40th	Median	60th	70th	80th	90th
Log of transport costs	0.0374**	0.0462***	0.0356***	0.0396***	0.0313***	0.0263***	0.0221***	0.0220***	0.0230***
	(0.0167)	(0.0111)	(0.00956)	(0.00836)	(0.00735)	(0.00612)	(0.00652)	(0.00633)	(0.00757)
Age	0.0106**	0.00855***	0.00738***	0.00760***	0.00764***	0.00999***	0.0123***	0.0136***	0.0131***
	(0.00426)	(0.00321)	(0.00249)	(0.00220)	(0.00199)	(0.00199)	(0.00227)	(0.00231)	(0.00256)
Male	0.115	0.239***	0.227***	0.209***	0.244***	0.261***	0.248***	0.241***	0.282***
	(0.0862)	(0.0618)	(0.0528)	(0.0446)	(0.0434)	(0.0384)	(0.0423)	(0.0445)	(0.0604)
Married	0.0635	0.0419	0.0600	0.0503	0.0503	0.0429	0.0811*	0.0785**	0.0979*
	(0.0811)	(0.0555)	(0.0430)	(0.0395)	(0.0393)	(0.0386)	(0.0424)	(0.0398)	(0.0508)
Race									
African									
Coloured	0.198	0.229**	0.181**	0.128*	0.161**	0.140**	0.212***	0.214***	0.250***
	(0.156)	(0.109)	(0.0821)	(0.0685)	(0.0630)	(0.0595)	(0.0628)	(0.0658)	(0.0811)
Asian/Indian	0.980***	0.614***	0.442***	0.236	0.288	0.556***	0.610***	0.456***	0.350**
	(0.236)	(0.167)	(0.153)	(0.174)	(0.218)	(0.212)	(0.167)	(0.147)	(0.151)
White	0.516***	0.602***	0.656***	0.631***	0.655***	0.638***	0.701***	0.604***	0.561***
	(0.177)	(0.133)	(0.103)	(0.0871)	(0.0871)	(0.0841)	(0.0886)	(0.0798)	(0.0931)
Highest level of education attained									
No schooling									
Primary	0.0635	0.0763	-0.0225	-0.0162	0.0183	0.00180	0.0310	0.0366	-0.0535
	(0.194)	(0.123)	(0.0937)	(0.0806)	(0.0773)	(0.0751)	(0.0845)	(0.0839)	(0.101)
Incomplete secondary	0.00419	0.0817	0.0522	0.0711	0.110	0.106	0.137	0.177*	0.142
	(0.178)	(0.130)	(0.101)	(0.0788)	(0.0706)	(0.0719)	(0.0868)	(0.0915)	(0.102)
Matric	0.295	0.263*	0.190	0.263***	0.315***	0.346***	0.393***	0.423***	0.384***
	(0.185)	(0.150)	(0.118)	(0.0897)	(0.0814)	(0.0798)	(0.0933)	(0.0983)	(0.108)
Certificate	0.0981	0.0875	0.108	0.129*	0.0887	0.112*	0.0594	0.0586	0.126
	(0.141)	(0.0959)	(0.0878)	(0.0706)	(0.0662)	(0.0635)	(0.0699)	(0.0737)	(0.0893)
Diploma	0.297**	0.368***	0.413***	0.366***	0.327***	0.352***	0.300***	0.285***	0.342***
	(0.145)	(0.0983)	(0.0806)	(0.0705)	(0.0716)	(0.0722)	(0.0768)	(0.0805)	(0.0906)
Degree	0.488**	0.568***	0.687***	0.641***	0.594***	0.589***	0.463***	0.498***	0.560***

	(0.201)	(0.149)	(0.119)	(0.0991)	(0.0998)	(0.0929)	(0.1000)	(0.108)	(0.122)
Household composition									
Household income per capita	0.101***	0.0379***	0.0341***	0.0343***	0.0290***	0.0181**	0.0220**	0.0290***	0.0218**
	(0.0188)	(0.0130)	(0.0111)	(0.00989)	(0.00886)	(0.00836)	(0.00918)	(0.00941)	(0.0109)
Main breadwinner	-0.121	-0.0151	-0.0216	-0.0160	0.00879	0.00490	-0.0411	-0.0374	-0.0697
	(0.0986)	(0.0691)	(0.0488)	(0.0400)	(0.0399)	(0.0373)	(0.0387)	(0.0436)	(0.0577)
Occupation category									
Elementary occupations									
Armed forces	0.981***	1.029***	1.224***	1.060***	0.933***	0.810***	0.754***	0.854***	0.811***
	(0.361)	(0.340)	(0.312)	(0.277)	(0.255)	(0.247)	(0.230)	(0.233)	(0.231)
Managers	0.764***	0.728***	0.895***	0.828***	0.764***	0.748***	0.887***	1.036***	0.996***
	(0.196)	(0.143)	(0.142)	(0.111)	(0.106)	(0.130)	(0.157)	(0.165)	(0.248)
Professionals	0.542***	0.544***	0.512***	0.516***	0.499***	0.494***	0.537***	0.607***	0.640***
	(0.160)	(0.119)	(0.0980)	(0.103)	(0.107)	(0.0916)	(0.0876)	(0.0881)	(0.103)
Technicians and associate professionals	0.0828	0.404**	0.500***	0.504***	0.450***	0.494***	0.582***	0.664***	0.747***
	(0.251)	(0.167)	(0.119)	(0.105)	(0.102)	(0.100)	(0.108)	(0.0984)	(0.138)
Clerical support workers	0.510**	0.485***	0.468***	0.447***	0.412***	0.357***	0.398***	0.424***	0.401***
	(0.245)	(0.128)	(0.100)	(0.0853)	(0.0816)	(0.0761)	(0.0820)	(0.0848)	(0.120)
Service and sales workers	0.153	0.0760	0.119	0.114*	0.0956*	0.0931	0.174***	0.160***	0.211***
	(0.119)	(0.0844)	(0.0733)	(0.0595)	(0.0576)	(0.0569)	(0.0609)	(0.0612)	(0.0766)
Skilled agricultural, forestry and fishery workers	-1.038	-0.729	-0.211	-0.187	-0.162	-0.185	-0.199	-0.319**	-0.276
	(2.240)	(1.063)	(0.480)	(0.314)	(0.229)	(0.192)	(0.154)	(0.154)	(0.201)
Craft and related trades workers	0.0606	0.0841	0.126	0.171**	0.131*	0.122*	0.126*	0.189**	0.198**
	(0.174)	(0.127)	(0.0988)	(0.0733)	(0.0685)	(0.0670)	(0.0748)	(0.0787)	(0.0981)
Plant and machine operators, and assemblers	0.171	0.179	0.237***	0.299***	0.232***	0.204***	0.236***	0.260***	0.153**
	(0.180)	(0.121)	(0.0797)	(0.0711)	(0.0658)	(0.0615)	(0.0627)	(0.0628)	(0.0665)
Industry									
Community, social and personal services									
Private households	-0.365**	-0.311***	-0.331***	-0.440***	-0.465***	-0.524***	-0.473***	-0.515***	-0.398***
	(0.162)	(0.106)	(0.0867)	(0.0790)	(0.0840)	(0.0784)	(0.0791)	(0.0752)	(0.0892)
Agriculture, hunting, forestry and fishing	-0.503***	-0.201	-0.264**	-0.354***	-0.381***	-0.394***	-0.401***	-0.399***	-0.222**

	(0.189)	(0.154)	(0.106)	(0.0929)	(0.0922)	(0.0829)	(0.0799)	(0.0776)	(0.106)
Mining and quarrying	0.370*	0.328**	0.464***	0.347***	0.358***	0.345***	0.347***	0.330***	0.284***
	(0.203)	(0.140)	(0.112)	(0.0968)	(0.0915)	(0.0866)	(0.0828)	(0.0830)	(0.105)
Manufacturing	-0.228	-0.217*	-0.139	-0.172*	-0.146*	-0.212***	-0.160**	-0.123*	-0.0125
	(0.176)	(0.116)	(0.0993)	(0.0877)	(0.0782)	(0.0701)	(0.0756)	(0.0745)	(0.0858)
Electricity, gas and water supply	0.238	0.126	0.0641	0.211	0.190	0.177	0.192	0.193	0.669
	(0.338)	(0.273)	(0.248)	(0.269)	(0.257)	(0.244)	(0.246)	(0.293)	(0.496)
Construction	-0.263	-0.142	-0.204*	-0.252*	-0.0854	-0.0750	-0.0297	-0.0308	0.0446
	(0.214)	(0.138)	(0.122)	(0.130)	(0.122)	(0.107)	(0.107)	(0.105)	(0.131)
Wholesale and retail trade; repair; hotels	-0.305**	-0.169*	-0.142**	-0.190***	-0.244***	-0.272***	-0.233***	-0.241***	-0.168**
	(0.132)	(0.0923)	(0.0715)	(0.0633)	(0.0658)	(0.0629)	(0.0661)	(0.0578)	(0.0744)
Transport, storage and communication	0.192	0.0723	0.146	0.0345	0.0397	0.0493	0.0537	-0.00937	0.0939
	(0.238)	(0.131)	(0.103)	(0.102)	(0.102)	(0.0879)	(0.0907)	(0.0975)	(0.100)
Financial intermediation, insurance, real estate	0.202	0.0424	0.0718	0.0101	-0.0487	-0.0737	-0.114*	-0.130*	-0.0608
	(0.131)	(0.0796)	(0.0804)	(0.0723)	(0.0714)	(0.0626)	(0.0653)	(0.0703)	(0.0877)
Union membership	0.265**	0.301***	0.246***	0.220***	0.230***	0.217***	0.237***	0.186***	0.171***
	(0.103)	(0.0690)	(0.0526)	(0.0446)	(0.0424)	(0.0412)	(0.0414)	(0.0409)	(0.0544)
Regional location of household									
Urban formal									
Rural formal	-0.130	-0.237**	-0.173**	-0.122*	-0.167***	-0.164***	-0.135**	-0.166***	-0.198***
	(0.136)	(0.106)	(0.0854)	(0.0663)	(0.0582)	(0.0534)	(0.0590)	(0.0574)	(0.0768)
Tribal authority areas	-0.172	-0.272***	-0.213***	-0.182***	-0.209***	-0.205***	-0.176***	-0.186***	-0.171***
	(0.120)	(0.0866)	(0.0729)	(0.0619)	(0.0567)	(0.0548)	(0.0593)	(0.0558)	(0.0648)
Urban informal	0.148	-0.0717	-0.0902	-0.102*	-0.175***	-0.147***	-0.130**	-0.127*	-0.0968
	(0.123)	(0.0805)	(0.0695)	(0.0544)	(0.0518)	(0.0539)	(0.0625)	(0.0722)	(0.0759)
Province									
KwaZulu-Natal									
Western Cape	0.152	0.156	0.189	0.169*	0.119	0.157**	0.109	0.00890	-0.101
	(0.203)	(0.144)	(0.122)	(0.0874)	(0.0726)	(0.0701)	(0.0801)	(0.0787)	(0.0889)
Eastern Cape	0.0948	-0.0936	-0.0872	-0.175**	-0.157**	-0.118*	-0.0858	-0.142*	-0.162*
	(0.142)	(0.126)	(0.0967)	(0.0851)	(0.0743)	(0.0693)	(0.0793)	(0.0733)	(0.0908)
Northern Cape	0.283	0.0392	0.0164	-0.0409	-0.0599	-0.0199	-0.0363	-0.0862	-0.106
	(0.189)	(0.154)	(0.122)	(0.0952)	(0.0859)	(0.0826)	(0.0902)	(0.0849)	(0.121)

Free State	0.0551	0.00306	0.0111	-0.0987	-0.122	-0.0980	-0.0428	0.00234	-0.000793
	(0.198)	(0.127)	(0.100)	(0.0825)	(0.0766)	(0.0764)	(0.102)	(0.0839)	(0.0976)
North West	0.371**	0.375***	0.245***	0.115	0.100	0.167**	0.117	0.0653	0.156
	(0.167)	(0.104)	(0.0915)	(0.0822)	(0.0708)	(0.0655)	(0.0716)	(0.0761)	(0.0999)
Gauteng	0.293**	0.254**	0.188**	0.0740	0.0925*	0.107*	0.106	0.0906	0.112
	(0.149)	(0.0996)	(0.0779)	(0.0640)	(0.0562)	(0.0546)	(0.0663)	(0.0654)	(0.0774)
Mpumalanga	0.313*	0.254**	0.174**	0.0910	0.0427	0.0578	0.0931	0.0913	0.228**
	(0.163)	(0.106)	(0.0841)	(0.0669)	(0.0600)	(0.0659)	(0.0764)	(0.0829)	(0.110)
Limpopo	0.121	0.210	0.133	-0.0256	0.0191	0.0368	0.0558	0.0688	0.0168
	(0.181)	(0.148)	(0.0993)	(0.0880)	(0.0842)	(0.0713)	(0.0801)	(0.0816)	(0.0919)
Constant	5.500***	6.022***	6.381***	6.624***	6.811***	6.885***	6.846***	6.966***	7.152***
	(0.305)	(0.240)	(0.192)	(0.155)	(0.136)	(0.139)	(0.153)	(0.155)	(0.173)
Observations	1,988	1,988	1,988	1,988	1,988	1,988	1,988	1,988	1,988

Notes: Standard errors are clustered at the household level and sampling weights are included. Strata with single sampling unit are centered at the overall mean.
Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table A5: Non-randomness of Non-reporting, 2010

	(1) Wage	(2) Transport Cost
Age	0.00916	-0.00928
	(0.0127)	(0.0144)
Male	-0.190	0.0336
	(0.255)	(0.278)
Married	0.203	-0.0586
	(0.235)	(0.303)
Race		
African/ Black	-	-
Colored	0.403	-0.302
	(0.440)	(0.871)
Asian/ Indian	1.209	1.364
	(1.141)	(1.240)
White	0.335	0.443
	(0.458)	(0.848)
Highest level of education attained		
No schooling	-	-
Primary	-0.451	0.106
	(0.457)	(0.699)
Incomplete secondary	-0.320	0.380
	(0.438)	(0.613)
Matric	0.0271	0.0929
	(0.457)	(0.621)
Certificate	-0.0615	0.359
	(0.447)	(0.438)
Diploma	-0.0838	-0.0715
	(0.316)	(0.470)
Degree	-0.0410	0.772
	(0.461)	(0.684)
Household composition		
Household income per capita	-1.80e-05	-0.000154**
	(3.29e-05)	(7.07e-05)
Main breadwinner	-0.214	-0.108
	(0.207)	(0.289)
Occupation category		
Elementary occupations	-	-
Armed forces	1.719**	
	(0.807)	
Managers	0.0984	-0.0636
	(0.566)	(0.697)
Professionals	0.250	0.278

	(0.529)	(0.590)
Technicians and associate professionals	0.522	0.353
	(0.415)	(0.543)
Clerical support workers	0.480	0.659
	(0.430)	(0.627)
Service and sales workers	-0.191	0.568
	(0.353)	(0.407)
Skilled agricultural, forestry and fishery workers	-1.480	
	(0.920)	
Craft and related trades workers	-0.258	-0.397
	(0.429)	(0.661)
Plant and machine operators, and assemblers	0.264	0.342
	(0.410)	(0.554)
Industry		
Community, social and personal services	-	-
Private households	-0.0290	-0.668
	(0.468)	(0.590)
Agriculture, hunting, forestry and fishing	-0.159	0.310
	(0.415)	(0.681)
Mining and quarrying	0.685	-0.769
	(0.480)	(0.999)
Manufacturing	-0.244	-0.582
	(0.497)	(0.472)
Electricity, gas and water supply	-1.007	0.132
	(0.834)	(1.025)
Construction	0.269	-1.026
	(0.513)	(0.745)
Wholesale and retail trade; repair; hotels	-0.232	-1.099**
	(0.403)	(0.470)
Transport, storage and communication	-1.437**	-0.901
	(0.631)	(0.777)
Financial intermediation, insurance, real estate	0.116	0.0561
	(0.430)	(0.424)
Union membership	-0.223	-1.305***
	(0.272)	(0.380)
Regional location of household		
Urban formal	-	-
Rural formal	0.168	-0.900
	(0.420)	(0.710)
Tribal authority areas	0.328	-0.0558
	(0.312)	(0.491)
Urban informal	-0.831	-0.551
	(0.526)	(0.467)
Province		
KwaZulu-Natal		
	-	-

Western Cape	0.721	0.268
	(0.643)	(0.878)
Eastern Cape	0.580	0.446
	(0.819)	(0.613)
Northern Cape	0.124	1.073
	(0.619)	(0.774)
Free State	0.797	0.126
	(0.613)	(0.644)
North West	0.346	0.626
	(0.563)	(0.840)
Gauteng	0.356	-0.333
	(0.633)	(0.576)
Mpumalanga	-0.0896	1.285**
	(0.742)	(0.574)
Limpopo	0.358	-0.302
	(0.574)	(0.775)
Constant	-3.116***	-2.149**
	(1.065)	(0.898)
Observations	3,411	2,280

Notes: Standard errors are clustered at the household level and sampling weights are included. Significance levels: *** p<0.01, ** p<0.05, * p<0.1

Table A6: Heckman Selection Estimation of the Effect of Transport Costs on Wages, 2010

Transport cost	0.0734*
	(0.0375)
Age	0.00519
	(0.00379)
Male	0.184***
	(0.0665)
Married	0.0979
	(0.0758)
Race	
African	
Colored	0.377***
	(0.102)
Asian/Indian	0.138
	(0.208)
White	0.552***
	(0.117)
Highest level of education attained	
No schooling	
Primary	0.134
	(0.219)
Incomplete secondary	0.266
	(0.202)
Matric	0.291
	(0.207)
Certificate	0.261***
	(0.0954)
Diploma	0.180**
	(0.0797)
Degree	0.595***
	(0.115)
Household composition	
Household income per capita	0.0316**
	(0.0156)
Main breadwinner	-0.0411
	(0.0703)
Occupation category	
Elementary occupations	
Armed forces	1.274***
	(0.224)
Managers	0.846***
	(0.134)
Professionals	0.379***
	(0.142)
Technicians and associate professionals	0.481***
	(0.142)

Clerical support workers	0.181
	(0.130)
Service and sales workers	0.0399
	(0.111)
Skilled agricultural, forestry and fishery workers	-0.403**
	(0.168)
Craft and related trades workers	0.0416
	(0.103)
Plant and machine operators, and assemblers	-0.0251
	(0.140)
Industry	
Community, social and personal services	
Private households	-0.277**
	(0.136)
Agriculture, hunting, forestry and fishing	0.142
	(0.179)
Mining and quarrying	-0.00963
	(0.186)
Manufacturing	-0.110
	(0.128)
Electricity, gas and water supply	0.345
	(0.427)
Construction	-0.0934
	(0.216)
Wholesale and retail trade; repair; hotels	-0.136
	(0.104)
Transport, storage and communication	-0.0963
	(0.165)
Financial intermediation, insurance, real estate	-0.169*
	(0.101)
Union membership	0.250***
	(0.0672)
Regional location of household	
Urban formal	
Rural formal	-0.283*
	(0.161)
Tribal authority areas	0.00711
	(0.126)
Urban informal	-0.00176
	(0.118)
Province	
KwaZulu-Natal	
Western Cape	0.0677
	(0.155)
Eastern Cape	-0.0224
	(0.164)

Northern Cape	-0.194
	(0.155)
Free State	0.0607
	(0.241)
North West	0.166
	(0.182)
Gauteng	0.0689
	(0.143)
Mpumalanga	0.154
	(0.152)
Limpopo	0.152
	(0.188)
Inverse mills ratio	-0.0522
	(0.0470)
Constant	6.759***
	(0.358)
Observations	958
R-squared	0.532

Notes: Standard errors are clustered at the household level and sampling weights are included. Strata with single sampling unit are centered at the overall mean. Significance levels: *** p<0.01, ** p<0.05, * p<0.1