



N.i.D.S.

NATIONAL INCOME DYNAMICS STUDY

Labour Market: Analysis of the NIDS Wave 1 Dataset

Discussion Paper no. 12

Vimal Ranchod
Southern African Labour & Development Research Unit
vimal.ranchhod@gmail.com

July 2009

1. Introduction

The purpose of this paper is to provide a brief summary of the labour market subset of the NIDS data. Most sections in this paper relate to the various sections in the adult questionnaire. We further include a short section on labour market information obtained from the proxy questionnaires, as well as the impact that this data has on the aggregate measures. We provide some personal commentary on our findings in each section, and highlight any findings or data related characteristics that may prove to be of interest to subsequent users of the data. All results are obtained using the post-stratification weights, unless specified otherwise.

The remainder of this paper is structured as follows: Section 2 provides a summary of the data at a broad level. Section 3 explores employment in more detail, including wage employment, self employment, casual employment and other types of non-wage employment. Section 4 describes working conditions, with an emphasis on hours worked and earnings. Section 5 relates to the unemployed, their experiences and expectations. Section 6 analyses search methods, both those used by people searching for jobs as well as those that were used by people who are currently employed. Section 7 considers the group who are not economically active. Section 8 incorporates information from the proxy questionnaires. Section 9 does a data quality consistency check by comparing summary statistics from NIDS with corresponding statistics obtained from the South African Labour Force Surveys conducted at approximately the same time. Section 10 considers the impact of various forms of interviewee non-response on the sample and Section 11 provides a concluding discussion

2. Aggregate trends and summary statistics

It is useful to have some context within which to place our labour market analysis. The NIDS data has 7 305 unique households, with a total of 28 255 household residents. Of these households, approximately 48 percent are in rural or tribal areas and 52 percent are in urban areas.¹ Mean household size in tribal areas is 4.6, while in urban areas the mean is 3.5. In the adult dataset, we have 15 639 respondents (aged fifteen or greater) who participated.

We categorize each adult into one of four mutually exclusive categories. 'Employed' is composed of people who are engaged in some type of productive activity, generally for the purpose of earning money. 'Searching unemployed' are people who are not employed, and have actively searched for employment in the past four weeks. 'Discouraged unemployed' are unemployed people who would have liked to have worked in the past four weeks, but have not actively searched for employment in that same time period. 'Not economically active' (NEA) are people who are not employed and do not want to find employment (for example, scholars/students, home-makers and the retired).

On aggregate, we have 6 012 employed, 1 838 searching unemployed, 976 discouraged unemployed and 6 663 NEA people in our data. There are also 150 respondents whom we cannot classify due to refusals or missing data for some of the relevant questions.²

In Table 1 below, we restrict our sample to adults aged 21 to 59 and apply the weights. We use this age group because 21 is the age of legal adulthood, and most people aged 20 or below are NEA. At age 60 or above, we find large amounts of retirement, due to pension age-eligibility. Overall, the percentage in each of the labour market categories are as follows: 55.2% are employed, 17.3% are actively searching, 5.9% are discouraged unemployed and 21.5% are NEA. A racial breakdown reveals that Africans are most likely to be NEA and least likely to be employed. They also have the highest unemployment rate, regardless of which definition is applied. By gender, females are twice as likely to be NEA, only about two thirds as likely to be employed, and have much higher unemployment rates.

If we look across the age cohorts, we observe some interesting, though well documented, phenomena. First, young South Africans struggle to find employment and have the highest

¹ These percentages correspond to the proportions in the sample, and exclude the effects of weights.

² The majority of these (128 observations) are clearly not employed, but have missing data on their preference for work or their recent search activity.

unemployment rates. These unemployment rates decrease monotonically with age. Second, labour force withdrawal occurs at relatively early ages in South Africa, with one third of 50 to 54 year olds in the NEA category. Third, even at the ‘peak’ employment years, in this case the people aged 35 to 44, fewer than two thirds of the adults are employed. Thus South Africa has relatively low employment rates in general, high youth unemployment, and early retirement. This life cycle trajectory has interesting implications when one considers household savings in conjunction with retirement planning and expectations.

Table 1: Summary statistics among prime aged adults

	# of obs.	Percentage in each category					
		i	ii		iv	v	vi
		NEA	Unemployed		Employed	Unemployment rate	
		Discouraged	Searching		(Strict)	(Broad)	
Aggregate	10120	21.5	5.9	17.3	55.2	23.9	29.6
African	7808	22.6	6.2	19.6	51.6	27.5	33.3
Coloured	1577	18.5	6.9	11.4	63.2	15.3	22.5
Indian	157	19.8	4.2	5.2	70.9	6.8	11.6
White	583	16.7	2.4	7.9	73.1	9.8	12.3
Male	3991	13.1	3.6	14.4	69.0	17.2	20.6
Female	6226	27.9	7.7	19.5	44.9	30.3	37.7
Age 20-24	1623	28.5	8.4	25.8	37.3	40.8	47.8
Age 25-29	1571	15.2	6.9	25.7	52.2	33.0	38.4
Age 30-34	1374	15.8	5.2	19.4	59.6	24.5	29.2
Age 35-39	1338	13.8	7.3	14.0	64.9	17.8	24.8
Age 40-44	1266	16.3	4.7	14.9	64.1	18.8	23.3
Age 45-49	1144	23.0	5.5	11.6	60.0	16.2	22.1
Age 50-54	1016	33.0	3.5	7.7	55.8	12.2	16.8
Age 55-59	885	44.0	2.5	2.7	50.8	5.1	9.3

Notes:

1. Sample restricted to adults aged 21 to 59 inclusive.
 2. All proportions have been weighted using the post-stratification weights.
 3. Unemployment rates are calculated using conventional definitions.
- Strict = (iii)/(iii+iv). Broad = (ii+iii)/(ii+iii+iv)

3. Employment

Table 2 provides the percentage within each different type of employment type, for various demographic groups. Of prime aged adults, 38.8% have a regular job, very few have a second job, 8.4% are self employed, 5.9% have had some casual employment in the past month and 4.4% have engaged in some form of subsistence farming. The percent who have helped in someone else’s business, possibly without payment, is 1.9%. Most people, if employed, have only one type of job.

Table 2: Mean percentage with various types of employment among prime aged adults

	Percentage in each category					
	i Regular Job	ii Second Job	iii Self Employed	iv Casual Work	v Subsistence Farming	vi Other Work
Aggregate	38.8	0.6	8.4	5.9	4.4	1.9
African	35.5	0.6	7.8	5.9	4.8	1.4
Coloured	48.8	0.7	4.7	7.6	0.9	3.5
Indian	44.9	0.0	16.5	9.1	12.1	3.7
White	55.4	0.6	14.4	3.4	2.2	4.0
Male	50.6	0.9	9.9	7.9	3.9	2.2
Female	29.9	0.4	7.4	4.5	4.8	1.7
Age 20-24	22.7	0.5	3.4	8.8	3.3	2.8
Age 25-29	37.6	0.6	7.8	5.1	2.4	1.6
Age 30-34	45.8	0.4	6.9	5.1	2.9	1.5
Age 35-39	47.7	0.3	11.5	5.6	4.3	1.4
Age 40-44	47.2	0.4	9.6	6.3	6.5	3.6
Age 45-49	40.3	1.2	9.6	6.6	6.6	2.1
Age 50-54	33.5	0.2	13.4	6.2	6.0	1.4
Age 55-59	33.2	1.8	9.3	3.0	7.6	1.0

Notes:

1. Sample restricted to adults aged 21 to 59 inclusive.
2. All proportions have been weighted using the post-stratification weights.
3. These categories are not mutually exclusive of each other.

African adults are the least likely to have a regular job, at 35.5%. Whites are the most likely to have one, at 55.4%. Indians and Whites are more likely to be self employed than the other groups. Whites, though, are least likely to engage in casual employment. Indians and Africans are more likely to engage in subsistence farming than Whites and Coloureds.³ Gender differences are fairly large. Men

³ The data with respect to Indians is imprecise due to the relatively small sample size.

are 20.7 percentage points more likely to have a regular job than women, 2.5 percentage points more likely to be self employed and 3.4 percentage points more likely to have had some casual work in the past month. Women are slightly more likely to have engaged in subsistence farming.

When comparing across the different age groups, we observe that youth are the least likely to have a regular job and the least likely to be self employed, and the most likely to have casual employment. The youth unemployment problem is thus likely to be compounded by high volatility in employment status. Self employment trends upwards with age, which may reflect capital constraints, skills constraints or network effects that change with time. Subsistence farming also increases with age. This might reflect cohort effects rather than aging effects, or the relative decrease in regular employment opportunities from the ages of 45 onwards.

3.1 Hours worked, Earnings and other working conditions

In this section we focus on the employed and their work conditions. Our analysis focuses on the relatively large employment categories, namely regular wage work, self employment, casual work and subsistence farming. Of these, regular wage work accounts for about 70 percent of employment.

3.1.1 Hours worked

In Table 3, we summarize the 'hours worked' distribution observed for the different groups of employed persons. We restrict the sample to individuals who reported 60 or fewer hours per week. This excludes slightly less than 10% of valid observations on aggregate.

Table 3: Distribution of weekly hours worked by employment category (%)

Hours	Regular Employment	Self Employment	Casual Employment	Subsistence Farming
0 - 4	0.6	6.9	32.4	27.3
5 - 9	6.5	18.5	28.2	24.4
10 - 14	3.2	13.5	15.8	18.4
15 - 19	1.1	4.1	7.0	10.8
20 - 24	1.3	7.6	4.0	7.2
25 - 29	1.1	3.9	3.8	3.5
30 - 34	3.1	6.6	1.3	3.9
35 - 39	6.4	4.0	2.5	1.0
40 - 44	37.5	14.3	1.7	1.8
45 - 49	29.3	13.3	2.3	1.3
50 - 54	6.9	5.8	1.0	0.4
55 - 59	3.2	1.6	0.2	0.0
# in category	3588	779	612	566
# with hours missing	284	61	80	50
# with hours >= 60	271	104	6	9
# with hours < 60	3317	675	606	557

Note:

1. *Sample restricted to adults aged 21 to 59 inclusive.*

2. *Percentages are weighted percentages, conditional on hours < 60.*

Among those with a regular job, the mass in the distribution occurs between 40 and 49 hours per week. This accounts for about two thirds of workers in this category. Approximately 10% work between 30 and 39 hours and between 50 and 59 hours per week, in each category. This yields a picture as expected with regards to the length of a conventional work week.

Among the self employed, we observe what seems to be a bimodal distribution. Of people who are self employed, 43% work fewer than 20 hours a week. A further 27.6% work a more conventional week of between 40 and 49 hours. Thus, self employment might coincide with a state of under-employment for some workers.

Casual employment, understandably, has a very different 'hours worked' distribution, with the mass of the distribution occurring at low numbers of hours. More than 80% of people who work in casual jobs work less than 20 hours per week. The same is true in subsistence agriculture. Thus, while these sources of employment are likely very important in alleviating poverty, the amount of work available may not be sufficient to fully empower people to escape from difficult economic circumstances.

3.1.2. Earnings

In Table 4, we analyze the distribution of net earnings in the various employment categories. One data concern that arises is that the number of people with missing data, generally due to refusals, is fairly high. If this is correlated with their wage levels, then the resulting sample will have a selection bias in it.

Of the regularly employed, 22% earn R1500 per month or less, while 52% take home at most R2500 per month. At the other end of the distribution, 20% earn more than R6000 per month. Of the self employed, 60.6% earn R1000 per month or less, although 20.7% earn more than R6000 per month, and 11% earn more than R16000 per month. This reinforces the idea that self employment, as a category, is comprised of a heterogeneous group of people with different experiences and payoffs. In the casual employment category, earnings are considerably lower. Almost two thirds or workers earn R1000 per month or less, while less than 3% earn more than R3500 per month. Subsistence agriculture is mostly an activity that does not generate any cash. 98.3% of prime aged adults that engage in this activity report that they received no cash at all from any produce.

The mean earnings amongst regular wage workers is R4284 per month. Some respondents only provided an income category rather than a specific value. We assumed a value equal to the midpoint of that category's range in such instances in order to calculate the mean. Among the self employed, the mean is slightly higher, at R5040 per month. This arises primarily due to the small proportion of fairly high earners in the R30000 to R50000 in the self employed group. Casual employment yields the least income per month, of these three groups, with a mean of just over R1000 per month. From an earnings perspective, it seems that the best type of employment to have is a regular job, unless one does not mind the high degree of dispersion in the earnings distribution that is observed among the self employed.

Table 4: Distribution of Earnings from employment (%)

Earnings	Regular emp.		Self Emp.		Casual Emp.		Subsistence Agric.	
	%	Cum.	%	Cum.	%	Cum.	%	Cum.
None	0.6	0.6	16.5	16.5	1.4	1.4	98.3	98.3
R1 - R200	0.7	1.4	13.5	30.0	16.1	17.4	0.6	98.8
R201 - R500	4.7	6.0	18.6	48.6	28.4	45.8	0.8	99.6
R501 - R1000	16.0	22.0	12.0	60.6	20.1	65.9	0.0	99.6
R1001 - R1500	13.4	35.4	2.5	63.1	15.4	81.3	0.0	99.6
R1501 - R2500	17.3	52.7	10.3	73.3	9.2	90.5	0.0	99.7
R2501 - R3500	11.2	63.9	1.6	74.9	6.6	97.1	0.3	100.0
R3501 - R4500	9.1	73.0	2.2	77.1	1.6	98.7	0.0	100.0
R4501 - R6000	7.0	80.0	2.3	79.3	0.6	99.3	0.0	100.0
R6001 - R8000	6.3	86.4	3.9	83.3	0.1	99.4	0.0	100.0
R8001 - R11 000	4.9	91.2	2.0	85.2	0.3	99.7	0.0	100.0
R11 001 - R16 000	5.0	96.2	3.8	89.0	0.3	100.0	0.0	100.0
R16 001 - R30 0000	2.7	98.8	3.4	92.4	0.0	100.0	0.0	100.0
R30 001 - R50 000	1.0	99.8	7.5	99.9	0.0	100.0	0.0	100.0
More than R50 000	0.2	100.0	0.1	100.0	0.0	100.0	0.0	100.0
Mean earnings	4284.0		5040.0		1002.0		18.7	
# in category	3,588		779		612		566	
# with missing income	983		350		173		15	
N	2,605		429		439		551	

Note:

1. Sample restricted to adults aged 21 to 59 inclusive.
2. Percentages are weighted percentages
3. Means are calculated using the midpoint of categories if only categorical data is available.
4. The overall mean of earnings is 3841.9 per month, using 4495 out of 5255 prime aged employed observations

3.1.3 Occupational Distribution

In Table 5 we look at the occupational distribution, and corresponding mean earnings, in the regular employment, self employment and casual employment categories. Of those in regular employment, there is full support for each occupational category, with the largest being elementary occupations, crafts and trades, professionals and service workers/salespersons. These make up 20.5, 15.7, 14.3 and 13.4% of the distribution respectively. Mean earnings are highest for legislators and professionals, at R12628 and R8263 in each category, and lowest among those in elementary occupations and agriculture and fisheries, with a mean of about R1330 in each group.

In the self employment group, over 50% are in elementary occupations, with mean earnings of R800 per month. A further 21.9% and 9.5% are in crafts or trades, and services or sales

respectively. Their corresponding mean earnings are R3378 and R3056 per month. Self employed professionals have the highest earnings, with a mean of R38232 per month.⁴ Of those in casual employment, the largest occupational groups are elementary occupations, services or sales, and crafts or trades, which account for 45.3, 15.6 and 15.1% of the distribution respectively. Mean earnings are highest among professionals, technicians and plant and machinery operators, at R1936, R1662 and R1656 per month. It is worth noting that elementary occupations and agriculture and fisheries comprise a large fraction of employment, while also having the lowest mean earnings in general.

Table 5: Occupational distribution and mean earnings

	Regular Employment		Self. Employment		Casual Employment	
	%	Earnings	%	Earnings	%	Earnings
Legislators	5.8	12628	2.7	2601	1.0	1122
Professionals	14.3	8263	6.1	38232	3.4	1936
Technicians	5.0	6576	0.7	50	1.2	1662
Clerks	11.0	4475	0.0	.	5.6	1532
Service workers/Sales	13.4	3297	9.5	3056	15.6	1003
Agriculture and Fishery	4.7	1334	0.5	18077	6.5	519
Crafts/Trade	15.7	3578	21.9	3378	15.1	1168
Plant and Machinery	9.6	2883	6.7	2687	6.3	1656
Elementary Occupations	20.5	1327	51.8	800	45.3	754
# of observations	3527	3050	273	166	569	512

Note:

1. Sample restricted to adults aged 21 to 59 inclusive.

2. Percentages are weighted percentages

3. Means are calculated using the midpoint of categories if only categorical data is available.

3.2 Job characteristics of regular wage workers

In this section, we briefly analyze some of the job characteristics of those in regular wage employment. In terms of remuneration, 36% report getting a 13th cheque, and 17% report some other type of bonus in the past year. These earnings would not have been captured in the earlier analysis of earnings.

⁴ This might be biased due to the midpoint imputation and the relatively few observations in this group. Nonetheless, the mean earnings are certainly much greater than the earnings of almost any other group, regardless of the type of employment they are in.

More than a quarter have deductions for medical aid, which might be subsidized by the employer. Almost half have pension fund deductions and close to two thirds have UIF deductions taken from their salaries. None of these would have been captured in the questions relating to net or ‘take-home’ pay discussed earlier. These proportions are not necessarily that high. A different viewpoint would be to observe that more than half of these workers do not have pension fund deductions, and over a third are not contributing to UIF. About a third of the workers belong to a trade union, and just over two thirds have an employment contract in writing.

Table 6: Job characteristics of regular wage workers

	%	Obs
13th Cheque	36.0	3539
Profit Share	3.7	3528
Other bonus	17.0	3521
Extra piece rate	3.4	3494
Medical Aid deduction	25.5	3528
Pension fund deduction	47.9	3531
UIF deduction	65.4	3533
Contract	67.3	3555
Trade union	33.6	3498

Note:

1. Sample restricted to adults aged 21 to 59 inclusive.

2. Percentages are weighted percentages

3.2.1 Unemployment

Unemployment is one of the most challenging socioeconomic problems in South Africa. Levels of unemployment are high, and have been for a very long time. A large proportion of the unemployed experience long spells of unemployment. Lack of employment correlates with poverty, and has implications for nutrition, schooling and inter-generational persistence. On the aggregate, this feeds into inequality and possibly social stability.

In this section, we investigate what information the data suggests about the unemployed. From Table 1, the unemployment rate among prime aged adults was 23.9 and 29.6 percentage points using the narrow and broad definitions respectively. Of these, 59.3 and 59.7 had never worked before. Among, the 21 to 29 age group, these percentages are 67.6 and 68.4 percent respectively. Of the discouraged unemployed, about half have been unemployed for three or more years. Of those actively searching, more than half have been unemployed for at least a year.

In Table 7, we show the distribution of reasons provided for the termination of the most recent job among the unemployed. By far the most common reason is that they got laid off, at about 60% using both definitions, which indicates a large amount of involuntary unemployment. The second most frequent reason is that the respondent was dissatisfied with their job, with more than 1 in 6 citing this as the reason. The unemployed might thus be composed of a complex mixture of voluntarily and involuntarily unemployed. Poor health is cited by close to 10% of respondents, and a non-trivial percentage is accounted for by the combined group of family and fertility related reasons.

Table 7: Reasons for leaving most recent employment

	Narrow	Broad
Health reasons	9.8	9.2
Caring for own children/relatives	1.7	2.4
Pregnancy	2.8	3.5
Other family/community responsibilities	1.31	1.6
Going to school	0.6	1.5
Lost job/job ended/laid off/business so	60.7	58.9
Changed residence	3.9	4.2
Dissatisfied with the job	18.1	17.8
Retired	1.0	0.9
N	370	542

Note:

1. Sample restricted to adults aged 21 to 59 inclusive.
2. Percentages are weighted percentages
3. Sample restricted to unemployed who have prior work experience

Fewer than 5% of unemployed respondents, using either definition, have rejected a job offer during their current spell of unemployment. This provides much stronger evidence that the unemployment problem is involuntary, and that a fundamental constraint is in the arrival rate of job offers. Given the NIDS evidence on the duration of unemployment, this is particularly worrying. The vast majority of the unemployed, if we trust the responses, have not received any job offers in more than a year.⁵

In Table 8 below, we summarize the expectations of the unemployed with regard to their chances of finding employment in particular time periods. The actively searching group is clearly more optimistic than the discouraged unemployed. Only half of the discouraged unemployed expect to

⁵ Of those who have rejected a job offer, more than 50% using the broad definition, and more than 60% using the strict definition, report having rejected the offer either due to low wages or not liking the job. However, this is unlikely to account for the aggregate unemployment patterns, since it applies to such a small proportion of the unemployed.

find unemployment within the next two years. Almost a quarter of the people actively searching for employment expect to find success within a month, and close to two thirds expect to do so within a year. When we combine these two groups, we find that only two thirds of the unemployed, on aggregate expect to find employment within the next two years. These beliefs seem consistent with the inferred data on job arrival rates mentioned above.

Table 8: Percentage who expect to find a job in given time intervals

Time Interval	Discouraged	Searching	Combined
0 - 1 month	11.1	23.5	20.4
0 - 6 months	26.5	44.6	40.0
0 - 1 year	44.2	63.6	58.7
0 - 2 years	50.4	71.4	66.1
Residual	100	100	100
N	786	1,584	2,370

Note:

1. Sample restricted to adults aged 21 to 59 inclusive.
2. Percentages are weighted percentages
3. The residual is comprised of people who do not expect to find a job within two years, and well as people with missing data.

Finally, we consider what levels of remuneration the unemployed consider to be reasonable. We compare various percentiles of this distribution with the corresponding percentiles of the observed distribution amongst regular wage workers.⁶ We do this for all prime aged adults as well as for youth aged from 21 to 29, in Table 9 below. This does not account for educational attainment, experience, special skills or any other unobservable factors that might affect a person's wages. Nonetheless, it provides a useful benchmark to determine whether the unemployed might be pricing themselves out of the labour market, particularly if one is willing to assume that the unemployed are likely to have less education, skills and/or experience.

⁶ We again use imputed values from the midpoint of categories if only categorical data is available.

Table 9: Comparing observed wages to 'reasonable' wages

Percentile	All prime aged adults		Youth aged 21 - 29	
	Observed	Reasonable	Observed	Reasonable
1%	200	500	200	500
5%	450	800	500	900
10%	700	1000	700	1000
25%	1100	1500	1000	1500
50%	2186	2000	1800	2000
75%	5000	3000	3500	3500
90%	9500	5000	6300	5000
95%	14000	6000	8600	7000
99%	33000	15000	13500	15000
Mean	4285	2783	2745	2987
N	3097	2208	772	1025

Note:

1. Sample restricted to adults aged 21 to 59, and youth aged 21 to 29 respectively.

2. Percentages are weighted percentages.

3. Observed wages are obtained from people in our dataset who have a regular job.

Where only categorical data was available, the midpoint of the category was used.

4. Both 'observed' and 'reasonable' expected wages are monthly take-home amounts.

When looking at all adults, we find that the lower percentiles of the expected wage distribution are higher than the observed distribution, but that this changes around the median. This pattern holds, broadly, amongst youth as well, although the cross-over point occurs around the 75th percentile. Based on this very simple diagnostic analysis, it might be the case that less skilled and/or experienced people in the unemployed group do have wage expectations that are slightly above what they are likely to obtain from employers.

4. Search Activities

We first look at the distribution of how people who are currently employed in a regular job found their jobs.⁷ This reflects some combination of the successfulness of the strategy, the frequency with which it was employed, as well as the job stability of the different types of jobs that particular search strategies are likely to yield.

The most common method is via social networks, with 40.5% finding their job through a friend or relative in a different household. A further 7.6% found their jobs through a household member, and 3.46 % through a previous employer. Thus, more than half of those currently employed in a regular job used some form of social capital to find employment. About 20% found employment by responding to advertisements, and about 15.7% found their job by 'cold calling' in person at factories, private households or shops.

Of those actively searching, the methods used to find employment vary, with about 28% of the searching unemployed using multiple search strategies simultaneously. The percentages are presented in Table 9 below. More than a third of job seekers employ a cold calling approach, 29.5% use social networks, and about 1 in 5 respond to advertisements. Other methods used including employment agencies, placing advertisements, using the internet and waiting at the side of the road.

Table 10: Frequency of search strategies employed

Search Method	%
Registered at an employment agency	14.3
Enquired at workplaces, farms, factories etc	36.9
Placed advertisement(s)	10.8
Answered advertisements	20.5
Searched through job advertisement(s) on the internet	8.5
Sought assistance from relatives or friends	29.5
Looked for land, building, equipment or applied for a permit	5.2
Waited on the side of the road	8.5
Sought financial assistance to start a business	0.9
Other	1.5

Note:

1. Sample restricted to adults aged 21 to 59 inclusive.

2. Percentages are weighted percentages

⁷ The sample considered is prime aged adults and the proportions are weighted using the post-stratification weights.

One outcome of having a low rate of success in finding a job is that the cost of searching might be prohibitive. In the dataset, the mean amount spent by the actively searching unemployed on transport costs related to the job search was R105.75 in the previous week, although 42.5% did not spend anything. About 1% spent R1000 or more. The main sources of funding for these transport costs were co-resident family members (67.8%), family members in other households (11.65%), savings (11.36%) and friends (6.79%).

5. The Not Economically Active group

From Table 1, the proportion in each group that is not economically active has a ‘U-shaped’ pattern as we consider successively older age groups. We present the main reasons for being NEA, by 10 year age groups, in Table 11 below.

Table 11: Reason for being Not Economically Active, by age group

Reason	Age group						All ages (20+)
	20 - 29	30 - 39	40 - 49	50 - 59	60 - 69	70 +	
I am too old	0.1	3.1	13.9	43.5	85.7	97.9	39.9
I am a full-time student	44.9	1.5	0.4	1.4	0.0	0.4	12.1
I am sick/disabled	7.2	26.5	37.3	34.3	11.4	1.0	17.4
I do not like the available jobs	2.4	3.7	6.1	0.8	0.3	0.0	2.0
I do not like working	2.8	1.8	1.7	0.1	0.1	0.0	1.2
I do domestic duties	4.5	13.4	10.5	6.9	0.1	0.3	5.2
I look after children	16.5	23.7	14.3	4.7	0.4	0.1	9.7
It costs too much to job hunt	7.4	4.4	1.5	1.9	0.2	0.0	3.0
The wages are too low	0.9	2.3	0.1	0.3	0.1	0.0	0.6
I spend my time cooking	2.4	2.7	5.9	4.1	0.6	0.1	2.4
Other	10.9	17.1	8.4	2.1	1.1	0.3	6.5
Total	100	100	100	100	100	100	100

Notes:

1. Percentages are weighted using post-stratification weights.

Amongst youth, the main reason for not being in the labour market is to pursue studies. This accounts for 44.9% of labour force withdrawal in the 20 to 29 age group, and 7.4% of youth are NEA due to the costs of searching for employment. Worth noting is the relatively high rates that both young and middle aged people report poor health as the main reason that they are not in the labour market. Older people, who make up the bulk of this group, are generally retired.

A combination of childcare and domestic duties is also important in explaining why young and prime aged adults are NEA. This accounts for 23.4, 39.8 and 30.6% of the 20-29, 30-39 and 40-49 year olds who are NEA, respectively. Almost all of these respondents are female, and the interaction between home production and market based production is likely to be important in understanding female labour force participation in South Africa.

We also have some information on why the NEA stopped searching for employment. Amongst prime aged adults, 29.6% report that they believed that searching for a job is pointless, a further 15.2% could not afford the costs of searching, 16% report pregnancy or the birth of a child and 11.4% report family responsibilities.

6. Labour market data from the proxy questionnaires

A final short section needs to be included because the adult sample is not fully representative of the adult population at large. People in residence halls, boarding schools, prison and hospitals would not have been able to participate in the survey. In addition, if a resident adult refused to participate, another household member might have been granted permission to provide information regarding that adult. These people were included by means of a proxy questionnaire. A majority of the proxies were obtained for the latter reason.

We include the proxy data with our adult dataset, and re-estimate the proportion in each labour market category among the prime aged adults.⁸ Table 12 presents the weighted percentages in each category in the adult dataset, proxy dataset and combined dataset.

Table 12: Proportion in each category, and effect of proxy data

	Adult data	Proxy data	Combined
NEA	21.5	14.7	20.9
Unemployed	23.3	24.0	23.3
Employed	55.2	61.3	55.8
N	10,120	1,138	11,258

Note:

- 1. Sample restricted to adults aged 21 to 59 inclusive.*
- 2. Percentages are weighted percentages*

The prime aged adults in the proxy questionnaire are more likely to be employed, and considerably less likely to be NEA. The effect on the aggregate distribution, however, is fairly insignificant since there are relatively few proxy adults.

⁸ The proxy data is limited in terms of the information contained. We cannot separate into searching and discouraged unemployed. Comparison to the adult dataset is done using the broad definition of unemployment.

7. Comparing to the SA Labour Force Survey.

We next compare the NIDS labour market data to corresponding data from Statistics South Africa's second quarterly Labour Force Survey of 2008. We chose this survey as the timing of the fieldwork overlaps considerably with that of NIDS, although they are not identical. The LFS covers the period from April to June, whereas the majority of the NIDS fieldwork was conducted from February to June.

We replicate Table 1 above using the LFS data, and then find the difference between the NIDS estimates and the LFS estimates. The relevant weights are used in all cases, and the samples are restricted to adults aged 21 to 59. We derived a variable from the LFS for labour market status, that would conceptually correspond as closely as possible to the variable we have from the NIDS data. Note that it is extremely unlikely that the differences will be zero, since the surveys had different questions, skip patterns, time periods, and involved a different sample.⁹

The differences in percentage points are presented in Table 13 below. The LFS is a much larger study, with more than four times as many respondents. The NIDS data has a 1.5 percentage point greater estimate of the number of NEA. It also has a higher estimate of the number of searching unemployed people, at 3 percentage points. The proportion employed in NIDS is considerably smaller, by 4.2 percentage points.

When we compare within racial groups, the NIDS data relatively underestimates African employment levels and the percent who are discouraged, and overestimates the percent who are searching and NEA. The largest discrepancy is for the African employed, who are 4.7 percentage points less likely to be employed in NIDS than the LFS. There are differences for each of the other race groups as well. Coloureds are less likely to be NEA and more likely to be discouraged unemployed, by 4.1 and 4.3 percentage points respectively. Indians have the largest absolute differences, but as mentioned earlier, the NIDS subsample of Indians is very small and likely to be imprecise for any inference specifically focused on this race group. Whites are less likely to be employed and more likely to be searching for employment, by 5.1 and 4.8 percentage points.

⁹ Note, however, that this only informs us as to how similar NIDS data is to the much larger and longer established LFS data. Any issues that are common to both, such as wealthier households being less willing to participate, will not generate any discrepancy here.

Differences in the proportions amongst males are relatively small. Amongst females, the NIDS dataset suggests lower levels of employment and higher levels of searching unemployment, by 6.1 and 4.8 percentage points respectively.

When we consider the different five year age cohorts, we find that NIDS consistently underestimates the proportion employed, and overestimates the proportion who are unemployed and searching. Amongst the older prime aged adults, it also overestimates the proportion in the NEA category.

Table 13: Difference between NIDS and SA LFS 2008 Q2 (in percentage points)

	# of obs. (in SA LFS)	Percentage in each category					
		i NEA	ii Unemployed Discouraged Searching		iv Employed	v Unemployment rate (Strict) (Broad)	
Aggregate	43,389	1.5	-0.2	3.0	-4.2	4.4	3.9
African	33,713	2.8	-1.3	3.2	-4.7	4.9	3.5
Coloured	5,057	-4.1	4.3	-1.4	1.2	-1.9	2.6
Indian	1,193	-8.3	2.9	-2.7	8.1	-4.4	-1.1
White	3,426	-0.9	1.2	4.8	-5.1	6.0	7.2
Male	19,366	0.2	-0.8	0.4	0.2	0.4	-0.3
Female	24,023	1.4	-0.1	4.8	-6.1	7.9	7.1
Age 20-24	6,805	-0.5	-0.7	2.4	-1.2	3.1	2.0
Age 25-29	7,198	0.0	-1.5	4.7	-3.2	5.5	3.8
Age 30-34	6,096	2.8	-1.7	3.2	-4.3	4.4	2.7
Age 35-39	5,857	-0.4	1.8	1.7	-3.2	2.5	4.0
Age 40-44	5,275	-0.3	0.6	5.7	-6.0	7.2	7.4
Age 45-49	4,737	2.9	1.5	4.3	-8.7	6.6	8.1
Age 50-54	4,120	6.4	0.2	2.8	-9.4	5.1	5.5
Age 55-59	3,301	5.3	0.3	-1.1	-4.4	-1.4	-0.6

Notes:

1. Sample restricted to adults aged 21 to 59 inclusive in both surveys.
2. All proportions have been weighted using the post-stratification weights from the relevant survey.
3. Unemployment rates are calculated using conventional definitions.
Strict = (iii)/(iii+iv). Broad = (ii+iii)/(ii+iii+iv)
4. Some differences arise due to different survey instruments and time frames.
5. We derived a corresponding 'state' variable to the one used in NIDS to maximize comparability.
6. The number of obs. refers to the LFS dataset, all other columns refer to the difference in the cell proportion, i.e. NIDS proportion less SA LFS proportion.
7. Some imprecision arises due to rounding errors at lower decimal values.

8. Interviewee non-response

An important issue to consider regarding how representative the NIDS data is, arises with regard to interviewee non-response. The sampling framework, in combination with the weights, should yield population level aggregates which are unbiased estimates of the national population at large. This can be challenged if there is widespread non-response, which is also correlated with other labour market variables of interest.

There are three levels at which non-response might occur. Firstly, the entire household may not have been interviewed. This could occur because in some cases it is impossible for the fieldwork team to contact a selected household. Alternatively, a household might refuse to participate in its entirety. Secondly, an individual in a household might refuse to participate. This might be particularly problematic if the employed are unavailable due to time constraints. The third level is that individuals may refuse to divulge some types of information which they feel is too sensitive or private. This often occurs when one requests information on income and assets. This too would bias inference of the national income distribution, and consequently, poverty and inequality measures.

The first level is partially corrected for by using the post stratification weights, although if the unavailability of the household is correlated with income, then this will reduce the representivity. The second level of non-response is also likely to generate problems. It seems probable that the individuals who place a high value on their time are more likely to refuse to participate, and that these are also likely to be people who are employed. Indeed, of the 1246 refusals in the adults questionnaire, 14% say they are “too busy”, 28.7% say it’s a “waste of time” and 14.7% were “Not home or unavailable”. This most likely leads to an underestimate of the employment levels.

Finally, there is non-response to particular questions. In terms of labour market status, this is not particularly acute, with only 150 observations which we cannot classify. However, with earnings this becomes much more problematic. In the regular employment, self employment and casual employment subsamples, 27.4%, 44.9% and 28.2% of respondents have missing values. This limits statistical power, and could induce significant bias in analyses that utilize this variable.

9. Conclusion

The new National Income Dynamics Study is an extremely exciting prospect for researchers and policy makers who are interested in issues relating to employment, unemployment, poverty and income inequality. At present, there is a wealth of data already in the Wave 1 questionnaires. Issues pertaining to earnings, the unemployed and labour force withdrawal are covered in some detail. The data also seem more or less in line with one would expect from a large national cross-sectional survey.

There exist some concerns regarding the data. These relate primarily to issues of non-response regarding questions about income, and non-participation among people who are likely to be wealthier. This affects the representivity of the data and in particular might lie behind the fact that NIDS seems to slightly underestimate employment rates relative to the LFS. Fortunately with a panel study, the fieldwork process for the second wave will contact all panel members whether or no they were successfully interviewed in person in Wave 1.

The full value of the project will only be realized with time, once the longitudinal structure can be used in analyses. However, even with the Wave 1 data, questions relating to the intergenerational transmission of opportunities, job stability, the interaction between fertility and the labour market, the impact of schooling on labour market experiences and several other important dynamic questions can be investigated in a way that is far better than has been possible in the South African context up to this point.