

University of Cape Town



CENTRE FOR
SOCIAL SCIENCE RESEARCH

**THE MYSTERY OF SOUTH AFRICA'S
GHOST WORKERS IN 1996:
MEASUREMENT AND Mismeasurement
IN THE MANUFACTURING CENSUS,
POPULATION CENSUS AND
OCTOBER HOUSEHOLD SURVEYS**

Martin Wittenberg

CSSR Working Paper No. 95



Published by the Centre for Social Science Research
University of Cape Town
2004

Copies of this publication may be obtained from:

The Administrative Officer
Centre for Social Science Research
University of Cape Town
Private Bag
Rondebosch, 7701
Tel: (021) 650 4656
Fax: (021) 650 4657
Email: kforbes@cssr.uct.ac.za

Price in Southern Africa (incl. VAT and postage): R 15.00

or it can be downloaded from our website
<http://www.cssr.uct.ac.za>

ISBN 1-77011-029-1

© Centre for Social Science Research, UCT, 2004

CENTRE FOR SOCIAL SCIENCE
RESEARCH

Southern Africa Labour and Development
Research Unit

THE MYSTERY OF SOUTH AFRICA'S
GHOST WORKERS IN 1996:

MEASUREMENT AND MISMEASUREMENT
IN THE MANUFACTURING CENSUS,
POPULATION CENSUS AND
OCTOBER HOUSEHOLD SURVEYS

Martin Wittenberg

CSSR Working Paper No. 95

December 2004

Martin Wittenberg is a Professor in the School of Economics at the University of Cape Town and a Research Associate at the Southern Africa Labour and Development Research Unit (SALDRU) within the Centre for Social Science Research (CSSR).

The Mystery of South Africa's Ghost Workers in 1996: Measurement and Mismeasurement in the Manufacturing Census, Population Census and October Household Surveys¹

Abstract

This paper compares estimates of total manufacturing employment from South Africa's 1996 manufacturing census, the 1996 population census and the October Household Surveys from 1995, 1996 and 1997. Findings show that there are 300 000 too few manufacturing workers recorded in the 1996 population census. Furthermore there are other inconsistencies between these data sources. Several possible explanations for the deficit in manufacturing workers are put forward, but none are compelling. It is concluded that aggregate employment series constructed from household surveys should be treated with caution.

"Is there any point to which you would wish to draw my attention?"
"To the curious incident of the dog in the night-time." "The dog did nothing in the night-time." "That was the curious incident," remarked Sherlock Holmes (*The Memoirs of Sherlock Holmes*, Silver Blaze).

Absences can be as telling as presences, as Sherlock Holmes reminds us. Some times, however, it is difficult to know whether one is really dealing with an absence or not. In the case of South African labour economics some absences have attracted attention: the surprisingly small size of the informal sector, or the surprisingly small rate of job creation during the 1990s. To these mysteries can be added another: the disappearance of about 300 000 manufacturing workers from the 1996 population census.

¹ The assistance of the Data First Resource Unit, University of Cape Town, and in particular of Lynn Woolfrey and Matthew Welch, is gratefully acknowledged.

1. Comparing the Manufacturing Census 1996 and the Population Census 1996

The substance of the problem can be seen from the last line of Table 1. The manufacturing census of 1996 reports a total of 1.4 million workers and an additional seven thousand “working proprietors”. By contrast the population census of the same year only records 1.1 million people employed in manufacturing.

Table 1 provides more geographical information about the mismatch. It is clear that some localised mismatches are probably due to commuting patterns. For instance statistical region 01 shows twenty-three thousand more jobs in manufacturing than are picked up in the population census. Ten thousand of these, however, live in the Kuils River magisterial district which falls into statistical region 02. That still leaves a surplus of over ten thousand manufacturing jobs in the greater Cape Town region. Similarly, the extra twenty thousand workers recorded by the population census in region 21 will be workers commuting to jobs in Pretoria. These still do not offset Pretoria’s total excess of fifty thousand.

These local deficits aside, in most cases the manufacturing census records substantially more manufacturing employment than the population census does. The biggest mismatches occur in Gauteng. The East Rand seems to have 78 000 more manufacturing jobs than workers and Johannesburg’s excess is about 58 000. The percentage difference (as percentage of the population census) is large in virtually all regions. What may account for these differences?

1.1 Explanation 1 - Measurement Error in the Manufacturing Census

Perhaps the information supplied in the manufacturing census was inaccurate. Is it plausible that employers would overstate their employment numbers? A number of factors would strongly militate against that assumption:

- a. In the context of increasingly stringent labour legislation (and attendant skills levies), it seems more plausible that employers would understate employment to quasi-official sources than overstate it.
- b. In an environment in which many firms began to outsource parts of their operations (due *inter alia* to changes in the labour legislation), one would expect firms to report on their “core” workforce and not the *de facto* workforce, some of whom would now consist of subcontracted employees.

Table 1 Manufacturing Employment in the Manufacturing Census and Population Census

<i>Subregion</i>	<i>Manufacturing Census 1996</i>			<i>Population census 1996</i>	<i>Difference</i>	<i>% difference</i>
	<i>Number of firms</i>	<i>Working proprietors</i>	<i>Number of workers</i>			
01 Cape Town	3,185	926	167,880	145,000	23,806	16%
02 Boland	648	180	32,232	43,911	-11,499	-26%
03 Southern Cape	113	50	3,336	4,444	-1,058	-24%
04 George, Mossel Bay	304	120	11,888	10,655	1,353	13%
05 Oudtshoorn	51	31	2,757	2,596	192	7%
06 Worcester, Ceres	174	73	8,379	9,906	-1,454	-15%
07 Saldanha	207	51	20,814	17,763	3,102	17%
08 West Coast	39	14	1,651	1,459	206	14%
09 Beaufort West	18	14	177	637	-446	-70%
10 Namaqualand	66	22	1,341	2,081	-718	-35%
11 Sutherland	9	4	66	284	-214	-75%
12 Victoria West	6	1	141	104	38	37%
13 Richmond, Colesberg	30	14	755	674	95	14%
14 Kuruman	52	21	1,114	1,036	99	10%
15 Carnarvon	12	11	105	222	-106	-48%
16 Herbert	42	20	1,303	1,206	117	10%
17 Kimberley	115	56	4,507	3,290	1,273	39%
18 Vryburg	34	14	999	2,185	-1,172	-54%
19 Lichtenburg	194	87	9,505	8,005	1,587	20%
20 Klerksdorp	142	66	4,812	3,658	1,220	33%
21 Brits, Odi	119	26	11,562	30,907	-19,319	-63%
22 Rustenburg, Madikwe	166	60	8,146	8,474	-268	-3%
23 Mmabatho	15	6	594	2,426	-1,826	-75%
24 Boshof	16	9	178	443	-256	-58%
25 Bloemfontein	263	72	14,790	13,868	994	7%
26 Smithfield	9	10	44	224	-170	-76%
27 Welkom	191	40	5,939	4,143	1,836	44%
28 Kroonstad	180	64	5,962	6,580	-554	-8%
29 Bethlehem, Harrismith	147	54	5,180	4,823	411	9%
30 Central Free State	60	53	1,683	3,106	-1,370	-44%
31 Sasolburg	55	8	7,585	5,000	2,593	52%
32 Witsieshoek	62	13	6,230	5,962	281	5%
34 Central Transkei	45	13	5,178	6,620	-1,429	-22%
35 Aliwal North	26	14	864	901	-23	-3%
37 Queenstown, Ciskei	72	23	3,256	2,995	284	9%
38 East London	409	86	36,901	29,964	7,023	23%
39 Albany	85	44	4,273	2,163	2,154	100%
40 Umzimkulu	0	0	0	710	-710	-100%
41 Cradock	21	15	397	530	-118	-22%
42 Humansdorp	55	25	2,378	2,058	345	17%
43 Port Elizabeth	780	223	65,919	47,917	18,225	38%

33, 36 & 44 Graaf-Reinet, Pondoland, Barkly East	31	26	700	3,546	-2,820	-80%
45 Steytlerville	5	6	584	144	446	310%
46 Ndwedwe	0	0	0	1,681	-1,681	-100%
47, 48 Durban/Pinetown	3,282	1,092	179,360	152,211	28,241	19%
49, 50 Nongoma	7	0	1,006	1,275	-269	-21%
51 Pietermaritzburg	431	202	26,538	19,239	7,501	39%
52 Camperdown	151	54	13,369	14,946	-1,523	-10%
53 Port Shepstone	160	92	7,104	5,010	2,186	44%
54 Underberg, East Griqualand	32	18	854	1,282	-410	-32%
55 Natal Midlands	109	47	8,127	6,642	1,532	23%
56 Ladysmith, Estcourt	171	34	22,349	13,277	9,106	69%
57 Newcastle	219	64	22,816	17,146	5,734	33%
58 Vryheid	48	14	2,191	2,977	-772	-26%
59 Richard's Bay	394	97	37,577	29,044	8,630	30%
60 Ubombo, Ingwavuma	13	7	1,688	2,133	-438	-21%
61 Highveld Ridge	80	25	12,548	9,349	3,224	34%
62 Witbank, Middelburg	206	57	17,579	15,012	2,624	17%
63 Groblersdal, Moutse	59	15	2,902	7,445	-4,528	-61%
64 Standerton, Ermelo	182	59	10,641	8,995	1,705	19%
65 Nelspruit	115	43	6,298	2,611	3,730	143%
66 Lowveld	183	52	20,441	16,007	4,486	28%
67 Phalaborwa	64	25	3,310	4,668	-1,333	-29%
68 Soutpansberg, Venda	118	76	5,527	7,327	-1,724	-24%
69 Potgietersrus	96	36	3,327	4,021	-658	-16%
70 Thabazimbi	28	8	572	1,275	-695	-55%
71 Sekhukhuneland	128	39	7,094	6,917	216	3%
72 Pietersburg	160	102	9,088	4,114	5,076	123%
73 East Rand	3,469	687	172,927	95,376	78,238	82%
74 Heidelberg, Springs	658	193	35,753	26,508	9,438	36%
75 West Rand	1,000	211	41,048	33,413	7,846	23%
76 Bronkhorstspuit	84	15	3,912	3,752	175	5%
77 Vereeniging, Vanderbijlpark	575	161	47,236	45,282	2,115	5%
78 Pretoria	1,405	283	93,736	43,000	51,019	119%
79 Johannesburg	3,908	727	141,279	83,554	58,452	70%
Total	25,788	7,200	1,420,302	1,128,109	299,393	

Notes:

1. Subregions correspond to the "Statistical Regions" used in the Manufacturing Census except for some in the North West province. The magisterial districts in the manufacturing census and the population census did not match up in this province (Wittenberg 2004). To get maximum comparability magisterial district 616 was incorporated with statistical region 22. This means that Rustenburg was excised from statistical region 21 and Koster from statistical region 19. Statistical region 47 was added to region 48, since Umlazi and Umbumbulu are functionally integrated with Durban. The other "joined" regions (e.g. 33, 36 and 44) were grouped together in the manufacturing census itself.
2. The population census data were extracted from the published "Community profile databases", Version 3. (Statistics South Africa, 1996)

- c. The business register (on which the manufacturing census was based) is widely acknowledged to have been out of date. It is least likely to have captured some of the small scale manufacturing firms which may have come into existence in the 1990s.
- d. The business register would not contain informal manufacturing enterprises. Informal sector employment is, however, covered by the census.

In short, one would expect the manufacturing census figures to be subject to considerable error: but in the opposite direction. One would expect the total employment estimate in the manufacturing census to be an **underestimate** rather than an overestimate. The mystery of the disappearing manufacturing workers in fact deepens.

1.2 Explanation 2 - Perhaps the Workers were not enumerated in the Census

If this were the case, it would be extremely troubling, since it would bear on the accuracy of all data sets at our disposal. It would be a failing of such colossal magnitude that labour economists would be inclined to commit collective *hara-kiri*.

There are perhaps two circumstances in which one might expect a manufacturing worker not to be enumerated:

- a. If the worker was living as a one person household in an urban area (e.g. in someone's back yard) and was not captured at census time.
- b. If the worker was an illegal immigrant and did not want to be enumerated.

The fact that the biggest mismatches recorded in Table 1 all happen in the big metropolitan areas may increase the likelihood that some of these factors may be at work. It is extremely unlikely, however, that these can explain the very large deficit. Firstly we know from other work that individuals with stable jobs are more likely to start families and become heads of households than other individuals. This should increase rather than decrease the probability of these individuals becoming captured in the census or in standard household surveys. Secondly, if a worker was in the country illegally, we would expect that person's employer not to report his or her presence either. Illegal workers can therefore not explain differences between the manufacturing census and the population census.

2. Comparing the Manufacturing Census 1996 to the October Household Surveys 1995-1997

If there was a systematic tendency to “miss” manufacturing workers at their place of residence, we would expect to see a similar mismatch between the manufacturing census and the October Household Surveys (OHS). As Table 2 shows, this is not the case. The bottom row of that table is more in line with the expectations outlined above. According to the 1995 and 1997 OHS, it appears as though total manufacturing employment should be of the order of 1.5 million. This would suggest that the manufacturing census missed about eighty thousand workers. The estimate from the 1996 OHS comes in under the manufacturing census total, but that survey was smaller than the typical OHS. Indeed the reason for reporting the estimates from the two years around 1996 is that there are concerns about how good the 1996 OHS could have been, given the demands of the census in that year.

In Table 2 we also report provincial totals. We observe that these also correspond much more closely with the provincial totals from the manufacturing census. The most notable exception is North-West, which has much lower recorded employment in the census than in any of the household surveys. The 1996 census total is not in any of the 95% confidence intervals for the household surveys. The reason for this may be cross-border commuting from the Winterveld area. In each survey, however, there are other mismatches which are less easy to rationalise. In 1995 the estimated level of manufacturing employment in the Eastern Cape is significantly below the 1996 manufacturing census figures. In 1996 it is Gauteng that is significantly below. In 1997 the Free State and Limpopo are significantly above the 1996 benchmark.

If one believes that even in the provincial breakdowns the manufacturing census is likely to give an underestimate, then relying exclusively on the household surveys one would have made serious errors in at least 1995 and 1996. The fundamental point is that estimating totals from survey evidence is likely to be a tricky business, given the fact that the estimates will be sensitive to the weights used. Furthermore given the clustered nature of the data, the standard errors around the totals are likely to be large. Indeed this can be seen in the 95% confidence intervals given in Table 2.

Table 2 Manufacturing Employment by Province in the Manufacturing Census and the October Household Surveys

Province	Manufacturing Census 1996	October Household Surveys											
		1995	95% confidence interval	In band?	1996	95% confidence interval	In band?	1997	95% confidence interval	In band?			
Western Cape	250,573	279,759	225,244	334,274	Yes	250,912	186,876	314,948	Yes	269,024	224,387	313,661	Yes
Eastern Cape	120,925	91,218	74,567	107,868	No	154,299	105,623	202,975	Yes	128,496	98,141	158,851	Yes
Northern Cape	9,481	12,203	6,174	18,231	Yes	15,669	5,315	26,023	Yes	15,407	9,253	21,561	Yes
Free State	47,914	42,612	32,646	52,578	Yes	54,036	34,391	73,681	Yes	67,740	50,425	85,055	No
KwaZulu Natal	324,700	336,865	283,799	389,931	Yes	295,505	220,035	370,975	Yes	333,660	286,700	380,620	Yes
North West	35,877	67,084	49,281	84,886	No	88,863	53,437	124,289	No	93,546	73,193	113,899	No
Gauteng	538,168	556,001	474,709	637,292	Yes	412,431	356,247	468,616	No	495,165	438,415	551,915	Yes
Mpumalanga	70,660	86,775	56,368	117,183	Yes	86,487	51,193	121,781	Yes	80,190	62,085	98,295	Yes
Limpopo	29,204	29,826	20,532	39,121	Yes	44,689	27,022	62,356	Yes	45,479	33,024	57,934	No
Total	1,427,502	1,502,342	1,324,661	1,680,023	Yes	1,402,891	1,181,226	1,624,556	Yes	1,528,707	1,444,715	1,612,699	Yes

Notes:

1. The figures for the 1995 OHS were calculated with 1996 weights.
2. The confidence intervals were calculated taking into consideration the cluster design of the surveys, but disregarding stratification.

Looking across the three surveys it is clear that not only are the confidence intervals large, but the point estimates jump around a lot. It is hard to believe that these are “real” changes. Given that the household surveys are supposedly designed to give reasonably accurate provincial breakdowns of variables, Table 2 would make one hesitant about placing too much weight on trends in the estimate of aggregates across different household surveys.

To return to the main topic of this paper, however, we note that there is no evidence to suggest that household based data gathering exercises would tend to miss manufacturing workers. If anything the data from the surveys would tend to suggest that the manufacturing census estimate of 1.42 million workers is probably an underestimate.

2.1 Explanation 3 – Misclassification of Employment Status in the Population Census

One possibility is that manufacturing workers may have been erroneously classified as unemployed or not economically active. *A priori* this seems unlikely. The manufacturing census is likely to capture mainly stable, formal sector jobs. These are the most likely to be reported in the population census. Indeed, the aggregate employment estimate from the population census (9.2 million) is in the same ballpark as the aggregates from the three household surveys (10 million, 9.2 million and 9.2 million respectively). One would expect that if 300 000 formal manufacturing jobs were reassigned to the unemployed that this would happen across the board in other sectors also, leading to substantially smaller estimates of total employment.

2.2 Explanation 4 - Misclassification of Industrial Sector in the Population Census

This is perhaps the most likely suspect. Indeed it turns out that there are 1.1 million individuals in the category “Industry not elsewhere classified or unspecified”. There is therefore ample room to believe that the disappearance of the manufacturing workers may be more apparent than real.

3. Comparing the Population Census 1996 to the October Household Surveys 1995-1997

A look at Table 3, however, starts to raise serious questions about this easy explanation. The first point to note is that it is not only manufacturing employment that is significantly below the point estimates emanating from the Household Surveys. Wholesale and Retail Trade and Services are also much lower. Indeed if we average the figures for the three household surveys and compare them with the population census, we note that there is a deficit of about 500 000 retail and wholesale workers and about 460 000 service workers. Given that the “residual” category is never less than 200 000, these two deficits are sufficient to explain the entire residual category!

Looking at the sectoral breakdowns, one is inclined to become much more sceptical about the prospects of doing any aggregate analysis on any of these data. The fact that mining employment in the population census is significantly higher than it is in any of the surveys speaks in favour of the census. It is well known that the surveys were badly designed to pick up mining employment (see, for instance, Klasen and Woolard 1999).

The fact that the census picks up much more employment in construction and in private households is worrying more for the quality of the household surveys than for that of the census. It is hard to see how manufacturing sector workers could erroneously have been placed into these categories. Instead one can envisage how the household surveys may have missed some domestic workers and small scale construction workers.

In particular years, there are other startling differences. The 1995 OHS picks up about four hundred thousand more agricultural workers than the census, or the other two household surveys, for that matter. It is hard to avoid the conclusion that the sampling frame for the 1995 OHS may have been less than perfect.

Similarly, the confidence intervals around the services sector figures for 1995 and 1997 do not overlap. One may conclude that employment in services collapsed between 1995 and 1997. With the increase in outsourcing by formal sector firms during that period, this would be surprising, to say the least.

Table 3 Employment by industry according to the Population Census 1996 and the October Household Surveys

Sector	Population Census 1996	October Household Surveys											
		1995	95% confidence interval		In band?	1996	95% confidence interval		In band?	1997	95% confidence interval		In band?
Agriculture	832,911	1,285,322	1,063,529	1,507,115	No	782,128	484,411	1,079,845	Yes	775,251	692,312	858,190	Yes
Mining	543,297	452,619	366,544	538,695	No	252,702	173,187	332,217	No	395,731	338,642	452,820	No
Manufacturing	1,128,109	1,502,342	1,324,661	1,680,023	No	1,402,891	1,181,226	1,624,556	No	1,528,707	1,444,715	1,612,699	No
Electricity	110,196	87,326	67,968	106,685	No	126,820	93,877	159,763	Yes	115,627	96,079	135,175	Yes
Construction	560,124	471,137	406,757	535,516	No	435,650	349,962	521,338	No	521,764	484,839	558,689	No
Trade	1,109,110	1,765,157	1,554,080	1,976,234	No	1,435,414	1,206,115	1,664,713	No	1,604,333	1,526,327	1,682,339	No
Transport	487,474	492,081	425,001	559,161	Yes	483,956	402,946	564,966	Yes	530,669	490,206	571,132	No
Finance	687,646	604,795	512,596	696,994	Yes	757,211	624,676	889,746	Yes	732,705	669,475	795,935	Yes
Services	1,594,542	2,247,757	2,009,593	2,485,921	No	2,030,902	1,722,554	2,339,250	No	1,894,277	1,802,105	1,986,449	No
Private households	1,065,052	847,394	735,598	959,190	No	817,770	621,756	1,013,784	No	765,973	716,006	815,940	No
Industry NEC		2,749	1,072	4,427									
Foreign govts	3,776												
NEC+Unspecified	1,108,484	200,624	166,840	234,407		625,892	528,063	723,721		360,743	326,023	395,463	
Total	9,230,721	9,959,302				9,151,336				9,225,780			

Note:

1. Industry classifications used in the October Household Surveys were the derived classifications supplied by Statistics South Africa.
2. Estimates were weighted using the person weights supplied by Statistics South Africa (1996 weights were used on the 1995 OHS).
3. Confidence intervals were based on standard errors corrected for clustering, but not stratification.

In short, the sectoral breakdowns do not support the simple hypothesis “Household Surveys Good, Population Census Bad”. It seems much more likely that there are strengths and weaknesses in both. The strength of the census is its coverage: there is much less room for doubt about whether the sampling frames captured particular types of individuals and households. One also does not have to be so concerned about whether the final results just happen to be driven by the particular sets of weights imposed by Statistics South Africa. The weakness of the census is that there would have been much less probing and checking of responses.

The key question in relation to our main puzzle is therefore whether one would expect misclassification of manufacturing workers to drive the “gap” that we observed between the population census and the manufacturing census. The first point we have already noted: we can explain virtually all of the “not elsewhere classified and unspecified” workers in the population census in terms of wholesale and retail trade and services. It seems much more plausible that given insufficient probing that these workers would be misclassified.

4. Comparing the Population Census 1996 to the Manufacturing Census (again)

A look at the provincial breakdowns given in Table 4 heightens these concerns. The biggest deficit in manufacturing employment census is in Gauteng. We would require fully 75% of everyone who was recorded under “industry not elsewhere classified and unspecified” in that province to have been a manufacturing worker, in order to get the population census and manufacturing census figures into alignment. In the other provinces, by contrast, we would require only small percentages of the residual category to be manufacturing workers. In short we would require processes of misclassification operating in Gauteng to be different to those operating in the other regions.

The simplest explanation that is congruent with these figures is the one advanced above: the residual category consists **largely** of individuals engaged in trade or services. It may perhaps contain a few misclassified manufacturing workers. It seems unlikely, however, that it will come close to making up the deficit of 300 000.

Table 4 Manufacturing Employment in the Manufacturing Census and Population Census by Province

<i>Province</i>	<i>Manufacturing Census</i>	<i>Population Census</i>	<i>Difference</i>	<i>Population Census: Industry NEC or unspecified</i>
Western Cape	250,573	236,371	14,202	130,096
Eastern Cape	120,925	97,548	23,377	128,254
Northern Cape	9,481	8,897	584	19,650
Free State	47,914	44,149	3,765	60,146
KwaZulu Natal	324,700	266,863	57,837	294,240
North-West	35,877	55,655	-19,778	61,990
Gauteng	538,168	330,885	207,283	277,193
Mpumalanga	70,660	59,419	11,241	62,068
Limpopo	29,204	28,322	882	74,847
	1,427,502	1,128,109		

4.1 Explanation 5 – The author’s data are corrupt and/or he is incapable of analysing them correctly

This explanation has much to recommend it. Indeed as Sherlock Holmes remarks, once all other possibilities have been eliminated, the one that remains has to be the truth. Regrettably this explanation would doom a few other analysts also. The sectoral figures given in Table 3 are similar to figures reported by Klasen and Woolard for 1995 (with different weights) and earlier years. The aggregate employment figures for the economy as a whole square with those given by Borhat (2003) and those for total manufacturing employment are very similar to figures reported by Borhat (2001). If there is an error, it is in the way the 1996 population census data have been extracted, courtesy of the SuperCross software.

There is also a major plot difficulty with dumping on the sleuth as a solution to the mystery. One couldn’t write this paper, for one. Even if explanation 5 is correct, it is not one that we could espouse.

Effectively this means that the mystery remains: where did at least three hundred thousand manufacturing workers disappear to in the 1996 population census? Some of them (maybe as much as a hundred and fifty thousand) may have been misclassified under “unspecified”. It still leaves a lot of unaccounted for

workers. Even more so, if one supposes that the true total should have been closer to 1.5 million.

How important is this mystery? My intuition is that it is extremely important. If we cannot reconcile the aggregates in one year, viz. 1996, we have absolutely no chance of talking meaningfully about changes in the aggregates from year to year. As it stands, we have three different pictures of what happened in that year: the manufacturing census, the population census, and the household survey. By comparing them we should get much more “traction” than we do for other years. The fact that we can’t reduce the uncertainty is troubling, to say the least. The evidence seems unambiguous that the population census number is far too low. It is much more difficult to pin down where those workers have gone to.

Conclusion

The disappearance of 300 000 workers would seem to be a mystery worthy of investigation by Sherlock Holmes himself. Unfortunately, this author is incapable of solving it. Nevertheless a number of useful insights have emerged from this investigation:

1. The Manufacturing Census figure of 1,427,502 employees and working proprietors seems like a realistic lower bound on manufacturing employment figures for 1996.
2. The sectoral aggregates for manufacturing, wholesale and retail trade and services recorded in the population census seem too low. The shortfalls in the latter two categories can probably be largely explained in terms of misclassification.
3. The sectoral totals estimated from the household surveys record too few mining workers and possibly too few construction workers and workers employed by private households.
4. The estimates of totals derived from the household surveys can be extremely noisy. It looks as though standard techniques may often give an **underestimate** of the true variability, even if one takes into consideration the clustered design. The additional variability possibly arises from:
 - a. Coverage problems (household surveys will tend to leave some sections of the population out).
 - b. Weighting issues.
 - c. Fieldwork errors and other types of nonsampling errors.

This is potentially an extremely important point, since many analysts wish to compare totals between years. Few of the exercises that I have seen have attempted to estimate standard errors for the totals.

5. Census information is potentially an extremely useful check on the survey data, particularly since one can do much more disaggregated analyses.
6. Being able to cross-check information from household survey type instruments against employer provided information (such as in the Manufacturing Census) can be extremely valuable. Regrettably the 1996 census of manufacturing is the most recent available data set.

We have identified the absence. Now we just need to find those missing workers....

References

- Bhorat, Haroon (2001), "Explaining Employment Trends in South Africa: 1993-1998", mimeo, Johannesburg: Friedrich Ebert Stiftung.
- (2003), "The Post-Apartheid Challenge: Labour Demand Trends in the South African Labour Market, 1995-1999", *DPRU Working Paper 03/82*, Cape Town: Development Policy Research Unit, University of Cape Town.
- Conan Doyle, Sir Arthur (1981) *The Penguin Complete Sherlock Holmes*, Harmondsworth: Penguin Books. First published 1930.
- Klasen, Stephan and Ingrid Woolard (1999) "Levels, Trends and Consistency of Employment and Unemployment Figures in South Africa", *Development Southern Africa*, 16: 3-35 .
- Statistics South Africa (1996a), *Population Census, 1996. Community profile databases*, Version 3. Report 03-01-30 (1996), Pretoria: Statistics South Africa.
- Statistics South Africa (1996b), *Census of Manufacturing 1996: Principal Statistics on a regional basis, Western Cape*, Report 30-01-04 (1996), Pretoria.
- *Census of Manufacturing 1996: Principal Statistics on a regional basis, Eastern Cape*, Report 30-01-05 (1996), Pretoria.
- *Census of Manufacturing 1996: Principal Statistics on a regional basis, Northern Cape*, Report 30-01-06 (1996), Pretoria
- *Census of Manufacturing 1996: Principal Statistics on a regional basis, Free State*, Report 30-01-07 (1996), Pretoria.
- , *Census of Manufacturing 1996: Principal Statistics on a regional basis, KwaZulu/Natal*, Report 30-01-08 (1996), Pretoria.
- *Census of Manufacturing 1996: Principal Statistics on a regional basis, North-West*, Report 30-01-09 (1996), Pretoria.
- *Census of Manufacturing 1996: Principal Statistics on a regional basis, Gauteng*, Report 30-01-10 (1996), Pretoria.

----- *Census of Manufacturing 1996: Principal Statistics on a regional basis, Mpumalanga*, Report 30-01-11 (1996), Pretoria.

----- *Census of Manufacturing 1996: Principal Statistics on a regional basis, Northern Province*, Report 30-01-12 (1996), Pretoria.

Wittenberg, Martin (2004) “The 1996 Manufacturing Census statistics imputed by municipality and district”, mimeo, Cape Town: School of Economics, University of Cape Town.

The Centre for Social Science Research

Working Paper Series

RECENT TITLES

Visser, M. Booysen, F. 2004. *Determinants of the Choice of Health Care Facility utilised by Individuals in HIV/AIDS-affected Households in the Free State Province of South Africa*. Cape Town. CSSR Working Paper No. 87.

Medupe, F. Collins, K. 2004. *Treatment for HIV/AIDS in the workplace: A Case Study of a Mine in Botswana*. Cape Town. CSSR Working Paper No. 88.

Mattes, R. 2004. *Voter information, government evaluations and party images, 1994-2004*. Cape Town. CSSR Working Paper No. 89.

Conradie, B. 2004. *Wages and wage elasticities for wine and table grapes in South Africa*. Cape Town. CSSR Working Paper No. 90.

Mattes, R. Mniki, N. 2004. *South Africa's Potential Brain Drain*. Cape Town. CSSR Working Paper No. 91.

Shin, D. Mattes, R. 2004. *Democracy as a lesser evil: Testing the Churchill notion of democracy in South Africa and South Korea*. Cape Town. CSSR Working Paper No. 92.

McCord, A. Ernst van Seventer, D. 2004. *The economy-wide impacts of the labour intensification of infrastructure expenditure in South Africa*. Cape Town. CSSR Working Paper No. 93.

Kahn, L. 2004. *Experiences of HIV/AIDS diagnosis, disclosure and stigma in an urban informal settlement in the Cape Peninsula: A qualitative exploration*. Cape Town. CSSR Working Paper No. 94.

The Centre for Social Science Research

The CSSR is an umbrella organisation comprising five units:

The Aids and Society Research Unit (ASRU) supports quantitative and qualitative research into the social and economic impact of the HIV pandemic in Southern Africa. Focus areas include: the economics of reducing mother to child transmission of HIV, the impact of HIV on firms and households; and psychological aspects of HIV infection and prevention. ASRU operates an outreach programme in Khayelitsha (the Memory Box Project) which provides training and counselling for HIV positive people

The Data First Resource Unit ('Data First') provides training and resources for research. Its main functions are: 1) to provide access to digital data resources and specialised published material; 2) to facilitate the collection, exchange and use of data sets on a collaborative basis; 3) to provide basic and advanced training in data analysis; 4) the ongoing development of a web site to disseminate data and research output.

The Democracy in Africa Research Unit (DARU) supports students and scholars who conduct systematic research in the following three areas: 1) public opinion and political culture in Africa and its role in democratisation and consolidation; 2) elections and voting in Africa; and 3) the impact of the HIV/AIDS pandemic on democratisation in Southern Africa. DARU has developed close working relationships with projects such as the Afrobarometer (a cross national survey of public opinion in fifteen African countries), the Comparative National Elections Project, and the Health Economics and AIDS Research Unit at the University of Natal.

The Social Surveys Unit (SSU) promotes critical analysis of the methodology, ethics and results of South African social science research. One core activity is the Cape Area Panel Study of young adults in Cape Town. This study follows 4800 young people as they move from school into the labour market and adulthood. The SSU is also planning a survey for 2004 on aspects of social capital, crime, and attitudes toward inequality.

The Southern Africa Labour and Development Research Unit (SALDRU) was established in 1975 as part of the School of Economics and joined the CSSR in 2002. SALDRU conducted the first national household survey in 1993 (the Project for Statistics on Living Standards and Development). More recently, SALDRU ran the Langeberg Integrated Family survey (1999) and the Khayelitsha/Mitchell's Plain Survey (2000). Current projects include research on public works programmes, poverty and inequality.
