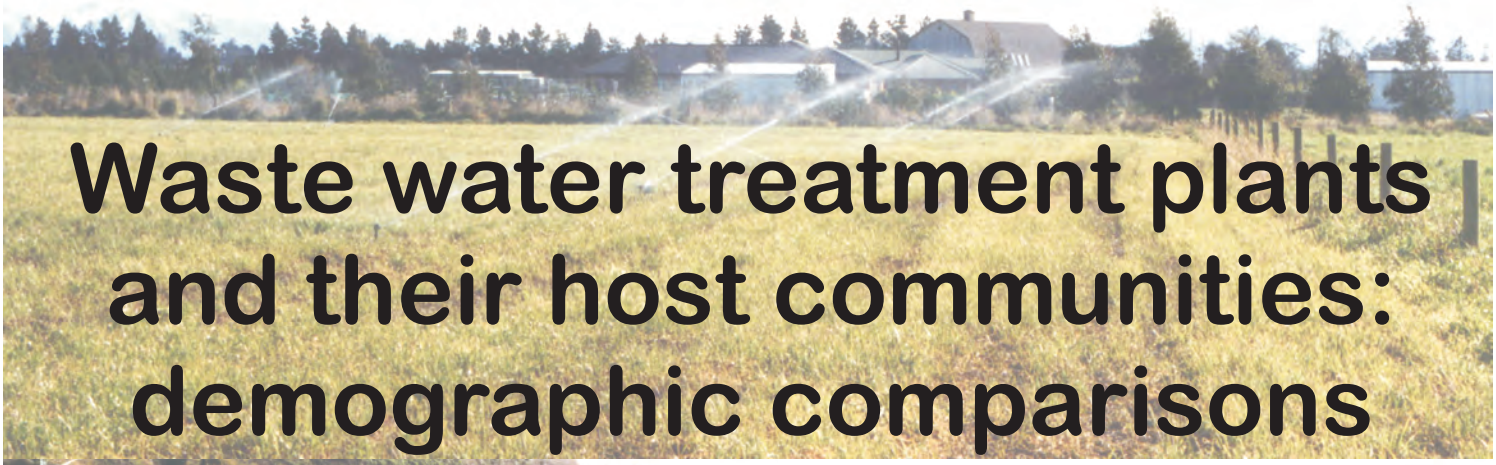
Host Communities: siting and effects of facilities



Waste water treatment plants and their host communities: demographic comparisons over time

James Baines

Taylor Baines & Associates

Host Communities: siting and effects of facilities

Waste water treatment plants and their host communities: demographic comparisons over time

By

James Baines

Taylor Baines & Associates

Working Paper FS21
Public Good Science Fund Contract TBA X0002

November 2001

Acknowledgements

The research team wishes to thank Gerard Fitzgerald for his professional advice concerning social indicators, Dr Patrick Graham for his professional advice on statistical methods, and Colin Goodrich for his critical review at various stages of the research.

Contents

LIST OF TABLES

LIST OF FIGURES

A: Introduction

The research reported here represents the third phase of investigations into the siting and effects of waste water treatment plant (WWTP) operations in relation to their host communities. Earlier phases of this research programme investigated:

Phase One: Historical patterns of WWTP siting from the perspective of the host communities involved (31 WWTPs) - reported in Working Paper FS2, Taylor Baines & Associates, 1999.

Phase Two: The day-to-day effects of WWTP operations experienced by members of the host communities, and community perceptions of the longer-term community impacts of these WWTP operations (eight case studies).

Phase One involved statistical comparisons of a specified set of demographic characteristics for host and source communities at the time the original siting decisions were taken. These historical comparisons applied to a sample of 31 WWTPs and their host and source communities. They were based on data sets drawn from the Census of Population and Dwellings carried out in 1976, 1981, 1986, 1991 and 1996. Similar historical comparisons were made between 'selected host communities' (i.e. host communities for sites ultimately selected) and 'short-listed alternative host communities' (i.e. alternative prospective host communities which were not ultimately selected). This latter comparison involved a sample of eight WWTP.

In both samples, the statistical comparisons were aimed at identifying whether or not there was any systematic bias towards relative disadvantage in host communities. As reported in Working Paper FS2 (ibid.), a comparison of host communities with source communities across the sample of 31cases revealed a consistent bias towards relatively disadvantaged host communities in New Zealand in the site selection process for WWTPs over the last twenty three years. The comparison of selected host communities with alternative candidate host communities revealed a similar bias, that during the process of site selection, more powerful candidate communities were consistently more effective in avoiding final selection.

This third phase of investigations is intended to provide an update on the statistical comparisons,made in Phase One, in order to determine whether or not any of the statistical differences identified previously have endured over time, or indeed, whether or not they have changed.

B: Methodology

Whereas the original historical analysis involved field research to investigate aspects of the community setting of selected sites and alternative sites, as well as the statistical comparisons of demographic characteristics for host and source communities, this analysis has been confined simply to the statistical component. The methodology is identical to that described in Section E of Working Paper FS2.

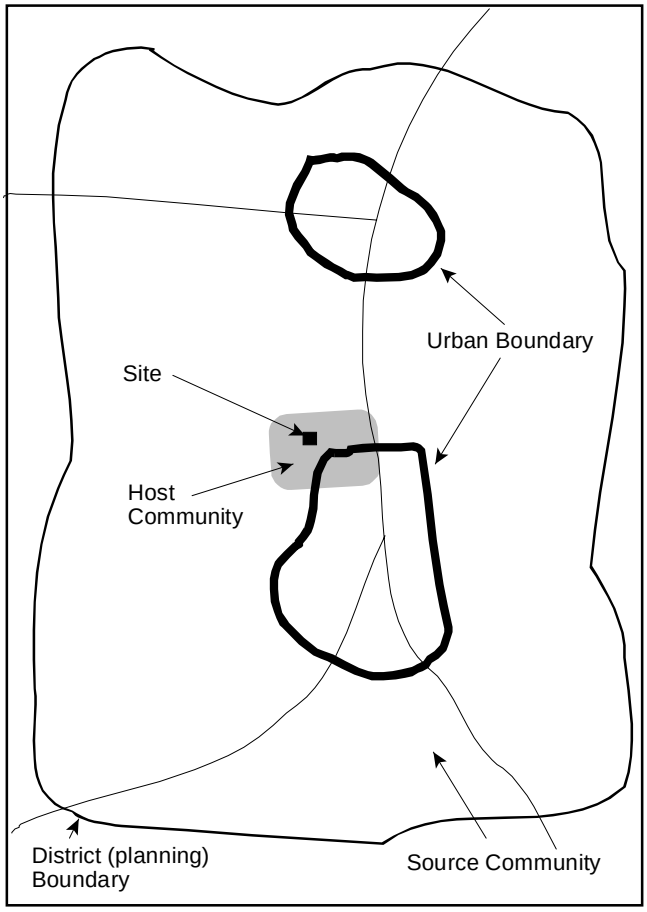
Definitions

As described on pp. 2-3 of Working Paper FS2, the research employs the concepts of '*source community*' and '*host community*'.

The term '*source community*' refers to the population resident in the geographic area that is serviced by the sewage reticulation network (see Figure 1).

The term '*host community*' refers to the community resident in the geographic area most clearly associated with a particular waste water treatment facility. It is not limited only to those residents who may experience direct physical effects from the operation of such a facility (such as the effects of odour or noise). It extends to include members of a coherent community - perceived in a social sense - that already exists at the time the facility siting decision was taken, and which will continue to develop in ways that may or may not be affected by the presence of the waste water facility in its locality.

Figure 1: Relating the concepts of 'host community' and 'source community'



Samples of facilities

The historical 'host-source' comparison reported in Working Paper FS2 involved a sample of 31 siting decisions. These included twenty two WWTPs which opened between 1975 and 1996, as well as nine sites for which the final site selection had been confirmed but which were yet to be opened. The historical 'host-alternative hosts' comparison involved a sample of eight WWTPs.

Since the most recent census data available are drawn from the 1996 Census of Population and Dwellings, there are no more recent demographic statistics to enable an update for any sites whose previous data also came from the 1996 Census. Results from the 2001 Census are not due for release until 2002. Thus the sample of sites remaining for the 'host-source' analysis in this report reduces to twenty two, while the sample for the corresponding 'host-alternative hosts' analysis in this report remains at eight.

Specifying indicators

As reported in Working Paper FS2 (p. 23 and Appendix I), in looking for appropriate indicators for this research, attention was focussed on socio-economic variables that could be considered to correlate with 'less powerful' groups in the community. 'Power' was defined as the ability to understand, take part in, or in any way influence the kind of decisions which lead to the siting of WWTPs, facilities which are usually subject to adverse expectations. This analysis takes its cue from research into environmental justice issues (e.g. Bullard, 1994). The seven variables[1] taken to correlate with powerlessness were:

- Educational qualifications — % of population aged 15 years and over with NO tertiary qualification
- Ethnicity — % of the total population that is NON European
- Occupational status — % of the population aged 15 years and over who are NOT legislators, administrators, professionals, technicians
- Employment status — % of the population aged 15 years and over who are NOT employed
- Life stage — % of households with at least one child under 5 years old
- Tenure — % of private dwellings that are NOT owned by the occupant
- Income — median household income[2]

Statistical method

The statistical method is identical to that reported in Working Paper FS2 (Appendix II). Host communities were compared with source communities by examining the ratio between 'observed' and 'expected' values[3] for each demographic variable and for each site. The statistics were pooled[4] for each variable over all sites to assess whether there is any indication of systematic bias in the demographic characteristics of selected sites. In other words, if the pooled statistics describing say the proportion of non-European residents or the proportion of unemployed residents for all the host communities appears different from the corresponding statistics for all source communities, then how strong is the statistical evidence for this difference?

For the reader's convenience, the original methodological note is reproduced in this report as Appendix I.

1 Based on the 'usually resident' population in the host and source communities.

2 This variable could not be subject to the same kind of significance testing as the others, being a median value rather than a bi-modal percentage value.

3 'Expected' values for a host community are derived by taking the corresponding value for the non-host community, assuming there were no difference between host and non-host values. Full details are provided in the methodological note in Appendix II.

4 The pooled statistics are based on weighted averages of the site specific results, with weights reflecting the amount of statistical information provided by each site.

C: Results

Summary of results from the previous historical analyses

The pooled results over all 31 WWTP and disposal sites revealed that host communities were typically different from source communities when analysed for indicators of 'power to influence decisions'. The results suggested a consistent bias towards relative disadvantage in host communities, as shown in Table 1 below[5]. Five out of six indicators suggested relative disadvantage with the strongest pooled indicators being ethnicity, life stage (children under 5 years in the household) and household tenure. This grouping of attributes is not especially characteristic of rural populations and is therefore likely to reflect the 'urban' components of host community.

Table 1: Pooled results for WWTPs (N=31)

	Education	Ethnicity	Occupation	Employment	Children	Tenure
Pooled Ratio Values	1.05	1.15	1.07	0.98	1.15	1.14
Host-source relativity	disadv.	disadv.	disadv.		disadv.	disadv.
Chi-square	153.78	215.38	315.00	0.63	47.58	198.17

Analysis of host-source comparisons for each individual site showed that some indicators of relative disadvantage were more consistent than others across the sample, as summarised in Table 2. This is most apparent for the indicators labelled 'Education', 'Occupation' and 'Children' where nearly every individual site ratio (R) is greater than unity. There was, nevertheless, considerable variation for some indicators within the sample of sites, as can be seen for the indicators labelled 'Ethnicity', 'Employment' and 'Tenure'. Nevertheless, the average R values shown in Table 2 corroborate the comparisons displayed in Table 1.

Table 2: Frequency of 'significant difference'

	Chi-square >3.84	Frequency	Average R
Education	R>1.00 R<1.00	11 sites 1 site	1.10
Ethnicity	R>1.00 R<1.00	11 sites 8 sites	1.18
Occupation	R>1.00 R<1.00	18 sites 1 site	1.14
Employment	R>1.00 R<1.00	3 sites 3 sites	0.94
Children	R>1.00 R<1.00	10 sites 3 sites	1.88
Tenure	R>1.00 R<1.00	12 sites 8 sites	1.43

[5] Where Chi-square values are less than 3.84, the ratios are not labelled to indicate relative disadvantage or relative advantage. This reflects the 5% statistical confidence convention - refer Appendix 1: Methodological Note.

Only descriptive analyses for the seventh variable - median household income - were undertaken since formal statistical analysis requires additional distributional information not provided routinely in census products. Analysis is therefore limited to simple ratios between the mean household income values of host and source communities, and to simple averages of these ratios across the sample and sub-samples of sites.

For the full sample of 30 sites[6], the average host/source ratio of median household income is 1.01, ranging from 0.69 to 1.65. Sixteen cases have a ratio <1.00 (i.e. the median household income for the host community is less than that for the source community) while 14 cases have a ratio >1.00. Thus, host and source communities are not clearly distinguishable on the grounds of household income distribution.

Table 3 shows the results of the analysis when the whole sample was sub-divided into two groups - siting decisions made under the Town and Country Planning Act (1977)(TCPA) and those made under the Resource Management Act (1991)(RMA).

Table 3: Pooled results for TCPA era (N=14) and RMA era (N=17)

	Education	**Ethnicity**	**Occupation**	**Employment**	**Children**	**Tenure**
TCPA: Pooled Ratio Values	1.05	1.24	1.07	1.02	1.17	1.14
Host-source relativity	disadv.	disadv.	disadv.		disadv.	disadv.
Chi-square	81.07	271.70	233.92	.039	28.32	39.07
	Education	**Ethnicity**	**Occupation**	**Employment**	**Children**	**Tenure**
RMA: Pooled Ratio Values	1.05	1.05	1.07	0.95	1.14	1.13
Host-source relativity	disadv.	disadv.	disadv.		disadv.	disadv.
Chi-square	72.97	12.03	81.46	2.56	19.65	148.87
Change from TCPA to RMA	NO change	MUCH LESS bias	NO change		LESS bias	NO change

In this analysis, one aspect stood out as having changed. What appeared to have been a much stronger ethnic bias under the TCPA regime, was much less apparent under the RMA. Life stage (children under 5 years in the household) appeared to have become slightly less prominent as well.

A similar trend was apparent from the comparison of median household income data, as shown in Table 4.

Table 4: Median household income ('mhi') ratios for TCPA era (N=13) and RMA era (N=17)

	Average 'mhi' ratio host/source	**range of 'mhi' ratios**	**distribution of 'mhi' ratios**
TCPA era (N=13)	0.90	0.69 - 1.27	R<1.00: 7 sites R>1.00: 6 sites
RMA era (N=17)	1.05	0.86 - 1.65	R<1.00: 9 sites R>1.00: 8 sites

[6] Median household income data was not available for one site opened in 1975, since neither the 1971 nor the 1976 census results provided this information.

Table 5 shows the results of analysing across two other sub-divisions of the sample - new sites versus upgraded sites.

Table 5: Pooled results for NEW sites (N=19) and UPGRADED sites (N=12)

	Education	**Ethnicity**	**Occupation**	**Employment**	**Children**	**Tenure**
NEW SITES:						
Pooled Ratio Values	1.06	1.04	1.07	0.95	1.06	1.17
Host-source relativity	disadv.	disadv.	disadv.		disadv.	disadv.
Chi-square	111.99	8.99	160.95	1.58	4.29	250.74
	Education	**Ethnicity**	**Occupation**	**Employment**	**Children**	**Tenure**
UPGRADED SITES:						
Pooled Ratio Values	1.04	1.37	1.08	1.00	1.25	0.92
Host-source relativity	disadv.	disadv.	disadv.		disadv.	disadv.
Chi-square	44.99	397.82	154.28	0.01	58.71	8.65

Three indicators varied markedly between these two sub-samples - ethnicity, life stage (households with children under 5) and household tenure. These are the same three indicators which showed the strongest bias when pooled across the entire sample of 31 sites (see Table 1 above). Non-europeans and families with children under 5 were much more prominent in the host communities of WWTPs which had been used for some time and were being upgraded. Furthermore, occupant ownership (rather than rental) was much more prominent in these communities. It could be that housing in the localities of WWTPs was more moderately priced, thereby encouraging greater levels of ownership amongst lower income households (non-european and young families). More detailed research would be required to establish whether or not this was the case. The longitudinal analysis reported here may assist in this, by showing whether or not site histories reveal characteristic demographic change after the establishment of WWTPs.

When the median household income data for these sub-samples (new sites versus upgraded sites) were analysed, little difference was evident (Table 6). The average host/source ratios were similar, both ranges spanned greater and less than unity, and individual ratios were reasonably evenly divided between greater and less than unity.

Table 6: Median household income ('mhi') ratios for NEW sites (N=18) and UPGRADED sites (N=12)

	Average 'mhi' ratio host/source	**range of 'mhi' ratios**	**distribution of 'mhi' ratios**
New sites (N=18)	1.00	0.69 - 1.65	R<1.00: 11 sites R>1.00: 7 sites
Upgraded sites (N=12)	1.03	0.86 - 1.34	R<1.00: 5 sites R>1.00: 7 sites

Table 3 indicates that non-europeans and families with children under 5 were less prominent in host communities created during the RMA era than the TCPA era. Table 5 indicates a similar contrast between new and upgraded sites. However, when these two factors are cross-tabulated - as in Table 7 below - it can be seen that upgraded sites were in fact more common in the sub-sample from

the RMA era than previously, which would suggest that both factors - a change in planning approach under the RMA, and the possibility that existing WWTP locations produced property price conditions which induced lower-income households into the area - may well have had an effect.

Table 7: Cross-tabulation of planning era (TCPA/RMA) with the incidence of new/ upgraded sites

	New	Upgraded	Total
TCPA	10	4	14
RMA	9	8	17
Total	19	12	31

The pooled analysis carried out for the eight 'host-alternative host' comparison revealed mixed results. Three demographic variables (education, ethnicity and household tenure) indicated a bias towards relatively disadvantaged candidate host communities in the short-list situation, while two showed no significant differences. As shown in Table 8, the strongest indications of social disadvantage (ratios >1.00) in the site selection process related to ethnicity and household tenure.

Table 8: Pooled results for comparing selected host communities with alternative host communities (N=8)

	Education	Ethnicity	Occupation	Employment	Children	Tenure
Pooled R	1.05	1.29	1.01	1.05	missing data	1.31
Host-alternative host relativity	disadv.	disadv.				disadv.
Chi-square	17.94	34.13	1.02	0.23		25.89

Analysis of individual sites showed considerable variation, as was the case for host-source community comparisons. Mixed results were also evident in the comparison of median household incomes between selected and alternative host communities. In five out of eight cases, the median household income level for the selected host community was lower than the alternative candidate host communities. The average ratio across all the seven sites for which data were available is 0.98 (ranging from 0.76 to 1.14), indicating that selected host communities have not been characteristically lower income communities than those which were not selected from the short lists.

Results of 'host-source' comparisons updated to 1996

To make a comparison over time, it was firstly necessary to repeat the base year calculations for the 22 WWTP or disposal sites which had opened prior to the 1996 census. Table 9 summarises the new pooled ratio values and compares them with the corresponding pooled ratio values for the full sample of 31 WWTP or disposal sites.

Table 9: Pooled results for 'host-source' base year[7] comparisons

	Education	Ethnicity	Occupation	Employment	Children	Tenure
Full sample of 31 sites:						
Pooled Ratio values	1.05	1.15	1.07	0.98	1.15	1.14
Host-source relativity	disadv.	disadv.	disadv.		disadv.	disadv.
Chi-square	153.78	215.38	315.00	0.63	47.58	198.17
Sample of 22 pre-1996 sites:						
Pooled Ratio values	1.05	1.14	1.07	0.97	1.10	1.63
Host-source relativity	disadv.	disadv.	disadv.		disadv.	disadv.
Chi-square	107.96	144.72	294.96	1.29	18.26	1464.97

The differences shown in Table 9 reflect the change in sample size - the reduction from a sample of 31 to a sample of 22. The overall pattern of host-source relativities is unchanged. However, reducing the sample has had a considerable affect on two variables. The greatest effect has been on the tenure variable, which shows a much higher incidence of rental housing in the vicinity of WWTP and effluent disposal sites[8] in the reduced sample of 22. A slight difference is shown for the life-stage variable (households with children under 5 years of age).

With the reduced sample, the average ratio of median household income[9] showed a very slight reduction - from 1.01 for the full sample of 31 sites, to 0.99 for the reduced sample of 22 sites.

Results from the base year analysis of this sample of 22 siting decisions have also been compared with the corresponding demographic data from the 1996 census - see Table 10 and Figure 2.

Table 10: Pooled results for 'host-source' 1996 comparisons

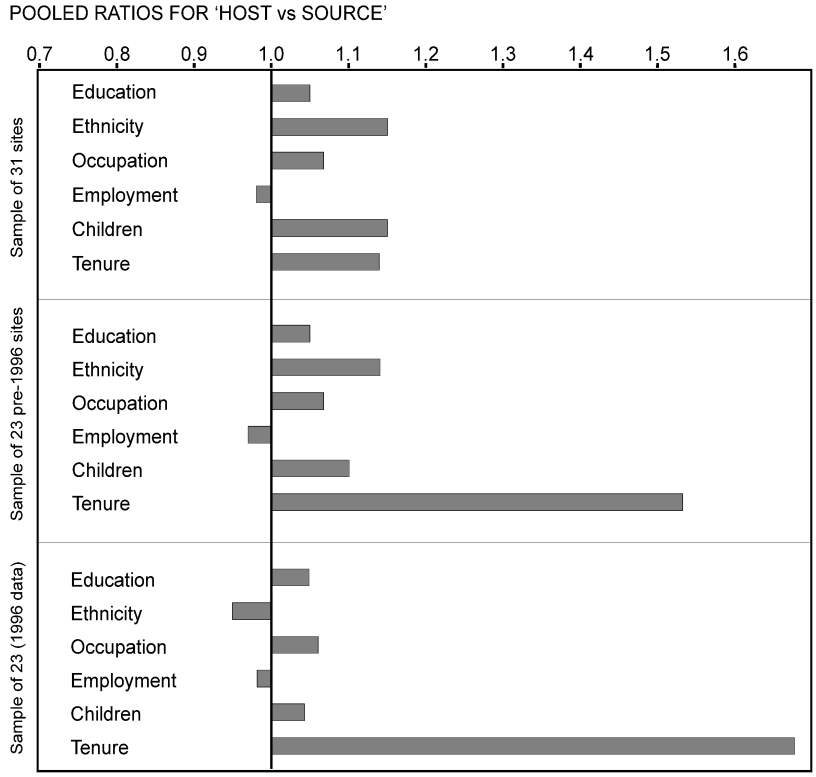
	Education	Ethnicity	Occupation	Employment	Children	Tenure
Full sample of 22 (pre-1996 data):						
Pooled Ratio values	1.05	1.14	1.07	0.97	1.10	1.63
Host-source relativity	disadv.	adv.	disadv.		disadv.	adv.
Chi-square	107.96	144.72	294.96	1.29	18.26	1464.97
Sample of 22 (1996 data):						
Pooled Ratio values	1.05	0.95	1.06	0.98	1.04	1.68
Host-source relativity	disadv.	disadv.	disadv.			disadv.
Chi-square	160.85	32.11	118.07	0.64	2.83	1573.10

7 'Base year' comparisons refer to data (prior to 1996) for site selection decisions.

8 Detailed analysis revealed that this change was due largely to removing a major metropolitan site from the sample. The affects of this on the other variables has been checked and confirmed.

9 If the ratio of median household income for host and source community is <1.00, household incomes are typically lower in the host community than the source community. An average of these ratios was calculated over the full sample of 31 sites and over the reduced sample of 22 sites.

Figure 2:



The results in Table 10 and Figure 2 reveal several changes in typical community composition for host communities with the passage of time. There has been a marked decrease (i.e. relatively disadvantaged status has improved markedly) in the proportion of non-European households in the vicinity of WWTPs. There is also no longer a significant difference in the proportion of households with young children living near these sites compared with the proportion of such households in the community at large (source communities). However, one pooled ratio (for tenure) has increased slightly (i.e. relatively disadvantaged status has further declined slightly) while three pooled ratios (for educational status, occupational status and employment status) have remained unchanged over time.

Table 11 presents the corresponding results for host-source comparisons of median household income. Recall that, for the simple ratios of median household income used in this analysis, ratios less than Unity (<1) indicate 'relative disadvantage' for the host community compared with the corresponding source community while ratios greater than Unity (>1) indicate 'relative advantage'.

Table 11: Pooled results of 'host-source' comparisons of median household income

	Median household income
Full sample of 22 (pre-1996 data):	
Pooled Ratio values	0.99
Host-source relativity	no difference
Sample of 22 (1996 data):	
Pooled Ratio values	1.04
Host-source relativity	slight adv.

The results in Table 11 show a trend that is consistent with those of the other demographic indicators.

Results of 'host-alternative host' comparisons updated to 1996

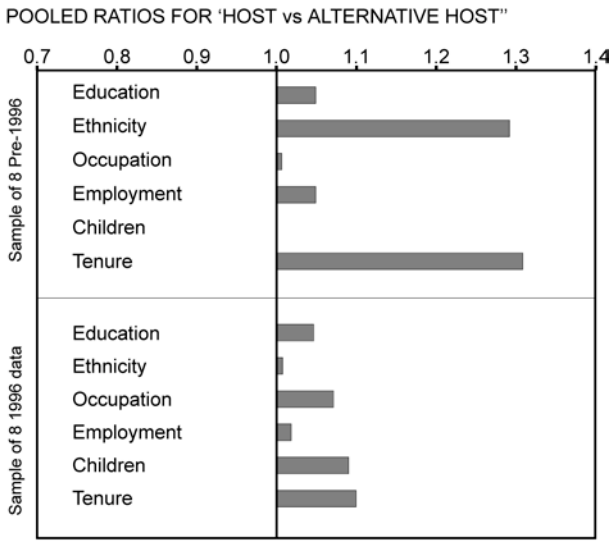
For this comparison there was no need to re-analyse the initial data set for a reduced sample of sites as all eight sites in the original comparison were selected prior to the 1996 census.

Results from the base year analysis of this sample of eight siting decisions will now be compared with results from comparing the corresponding demographic data from the 1996 census - see Table 12 and Figure 3.

Table 12: Pooled results for 'host-alternative host' 1996 comparisons

	Education	Ethnicity	Occupation	Employment	Children	Tenure
Full sample of 8 (pre-1996 data):						
Pooled Ratio Values	1.05	1.29	1.01	1.05	missing data	1.31
Host-source relativity	disadv.	disadv.				adv.
Chi-square	17.94	34.13	1.02	0.23		25.89
Sample of 8 (1996 data):						
Pooled Ratio Values	1.05	1.01	1.07	1.02	1.09	1.10
Host-source relativity	disadv.		disadv.			disadv.
Chi-square	20.01	0.08	24.59	0.15	3.52	4.82

Figure 3:



The results in Table 12 and Figure 3 reveal that with the passage of time, pooled ratios for ethnicity and rental accommodation have decreased markedly (i.e. relatively disadvantaged status has improved markedly) while occupational status has emerged as a symptom of relative disadvantage. No change was observed for indicators of educational status and employment status, and no trend

could be evaluated for the life-stage variable (households with young children) because of the absence of earlier census data.

Table 13 presents the corresponding results for host-alternative host comparisons of median household income.

Table 13: Pooled results of 'host-alternative host' comparisons of median household income

	Median household income
Full sample of 8 sites with alternatives (pre-1996 data):	
Pooled Ratio values	0.98
Host-alt.source relativity	no difference
Sample of 8 sites with alternatives (1996 data):	
Pooled Ratio values	1.04
Host-alt.source relativity	slight adv.

The results in Table 13 show a trend that is consistent with those of the other demographic indicators - any relative disadvantage for these eight cases has not become more accentuated over time; indeed, a slight improvement in relative income status appears to have occurred.

D: Discussion and Conclusions

Basis for interpretation

The quantitative statistical comparisons presented in this Working Paper are based on a group of discrete demographic indicators which reflect the comparative status of pairs of communities with respect to educational status, ethnicity, occupational status, employment status, the prevalence of families with young children, and the status of dwelling tenure.

Originally, these demographic indicators were selected because they were considered to correlate with a community's ability to participate in and influence the planning process. Specifically, the research inferred that communities with a population prevalence of low educational achievement, minority ethnic groups, low occupational status, high unemployment rate, households with young children, low dwelling ownership, and low median household income would be likely to have greater difficulty in understanding, participating in, or in any way influencing the kind of decisions which lead to the siting of WWTPs. It was intended that this inference be taken in a comparative rather than an absolute sense, which is particularly relevant to cases where the planning process involves consideration of several alternative or 'competing' prospective sites.

In the context of this longitudinal analysis, the same grouping of deomographic indicators can be taken to correlate with a community's ability to participate in and influence on-going management of the WWTPs, particularly when Territorial Local Authorities have generally not been pro-active in establishing mechanisms for community involvement.

As was the case with the results presented previously in Working Paper FS2, the emphasis is on the pooled comparisons across samples of sites and their communities. It should also be added that the primary emphasis for interpreting these results should be on the cluster of indicators taken as group rather than individually. A single demographic indicator, considered in isolation, may not be indicating 'powerlessness'. It is the compounding effect of a group of indicators - particularly when a consistent pattern emerges - that is likely to indicate a condition of relative powerlessness.

Trends identified

The results of both longitudinal analyses (i.e. 'host-source' comparisons and 'host-alternative host' comparisons) point to the possibility of an interesting social and cultural phenomenon - that Maori households may have elected to move away from the vicinity of WWTPs and their effluent disposal sites. This apparent trend could also be linked to some of the other apparent changes - a reduction in households with young children, and an increase in households with low occupational status. However, such possibilities will require further targeted research to confirm. Whether or not the apparent migration of non-European households can explain all the changes in the other variables cannot be deduced without further research.

Updating the comparison of host and source communities reveals a progressive change over time from five indicators of relative disadvantage to three. Apart from the reduction in non-European households and the reduction in households having young children, there appears to have been a slight increase in the incidence of host community households occupying rental accommodation. It might be assumed that this reflects similar trends in the actual incidence of rental accommodation in

the vicinity of WWTPs and effluent disposal sites, resulting perhaps from people buying in a location but choosing not to live there, or people owning, not being able to sell, and therefore renting to provide some return. However, this trend would run counter to the comparison reported previously (Taylor Baines & Associates, 1999) where the host communities around upgraded WWTPs[10] exhibited significantly lower proportions of households renting than host communities around new sites (refer to Table 5). This is a question that will need follow-up research.

Updating the comparison of host versus alternative host communities reveals that the main indicators of relative disadvantage at the time of site selection - ethnicity and tenure - have reduced significantly. The preponderance of non-European households in host communities has disappeared altogether, a trend which is possibly related to the appearance of relative disadvantage in occupational status. However, this will also need further research.

Overall, the results of both longitudinal analyses suggest that host communities have not systematically become more disadvantaged as a result of hosting WWTPs. In fact, the differences evident at the time of site selection appear to have diminished.

Conclusions

This report has demonstrated that some changes in demographic comparison for host communities of WWTPs have occurred over time. Whether or not these relative changes (in demographic comparisons) are in any way attributable to the way the facilities have been operated and managed cannot be deduced from this analysis. It is possible, even likely, that other factors for change have had a bearing over time. It is likely that neighbours become accustomed to effects over time. They may also become more realistic and accepting - it was a common experience in the case studies carried out that other activities in the locality of a WWTP do produce similar kinds of effects, particularly odours, but with less negative connotations (e.g. pigs farms, chicken farms, dairy milking sheds, and so on). Furthermore, another finding of the case study research was the positive experience of improved water quality or the positive landscape and wildlife changes around some wetland developments. In a minority of cases, actual experience of effects and impacts has proven to be less intrusive than expected.

All these factors have the possibility of contributing to a less negative perception about the neighbourhood of a WWTP.

With these factors in mind, this analysis does suggest that the siting of WWTPs has not resulted in some kind of long-term social and community blight, when assessed using the indicators employed in this piece of social research.

This analysis also points to several areas of research into social-environmental interactions and community change requiring more focussed attention - the apparent response of non-European households, and the dynamics of dwelling ownership and owner-occupancy around NIMBY facilities as the host community gains experience of actual effects.

[10] i.e. those host communities with previous experience of WWTP operations.

References

Taylor Baines & Associates (1999): Waste Water Treatment Plants and their host communities. Working Paper FS2.

Taylor Baines & Associates (2000): Contemporary host community experience of waste water treatment plants in New Zealand. Working Paper FS19.

Bullard, Robert D (1994): "Dumping in Dixie: Race, Class and Environmental Quality" Boulder Colorado: Westview Press.

Appendix I: Methodological note

Notes on statistical analysis for Siting of Landfills project

Patrick Graham

pgraham@chmeds.ac.nz

25th June 1998

1 Preliminary comments

1.1 estimation or hypothesis testing

Given that census data is being used for the comparisons, it is likely that sample sizes will be large and that classical significance tests will yield "significance" even for differences which are fairly trivial. It may be better to focus on estimation of differences, via confidence intervals, rather than hypothesis testing.

1.2 Types of comparisons

Comparisons can either be made between host areas and source areas or between host and non-host areas and both relative and absolute differences

between host and comparison areas can be considered.

1.3 Notation

p_h –proportion with the characteristic (e.g.) tertiary qualified in the host area.

p_{nh} – proportion with the characteristic in non-host areas.

N_h – number of individuals resident in the host area.

N_{nh} – number of individual resident in the non-host area.

N_s – number of individuals resident in the source area.

$p_s = (N_h p_h + N_{nh} p_{nh})/N_s$ – proportion with the characteristic in the source area.

2 Statistical methods for comparing host areas with source areas.

2.1 inference for observed:expected ratio

The ratio of the proportion with the characteristic in the host area to the proportion expected if the overall source area prevalence of the characteristic held also in the host area is $R = p_h/p_s$. Suppose $R = 1.3$ and the characteristic of interest is the proportion renting their home, then we are saying that the proportion of people renting is 30% greater in the host area than in the source area as a whole. Note this is *not* a difference of 30 percentage

points, i.e. we are talking about the difference between, say, 10% and 13% not 10% and 40% In order to develop methods for statistical inference for the observed: expected ratio it is easiest to work on a log scale (natural logarithm) and convert back to the original scale at the end. This is illustrated below.

The variance of the natural logarithm of the observed:expected ratio, R, is

$$V(\ln(R)) = \frac{V(p_h)}{p_h^2} - 2\frac{Cov(p_h, p_s)}{p_h p_s} + \frac{V(p_s)}{p_s^2}$$

where

$$V(p_h) = \frac{p_h(1-p_h)}{N_h}$$

$$Cov(p_h, p_s) = \frac{p_h(1-p_h)}{N_s}$$

$$V(p_s) = \frac{n_h p_h(1-p_h) + N_{nh} p_{nh}(1-p_{nh})}{N_s^2}.$$

A 95% confidence interval for the observed to expected ratio is obtained as

$$\exp(ln(R) \pm 1.96\sqrt{V(ln(R))})$$

A significance test can be based on the statistic

$$\chi^2 = \frac{(ln(R))^2}{V(ln(R))}$$

which can be compared to the chi-square distribution with 1 degree of freedom. Values of χ^2 greater than 3.84 are said to be statistically significant at the 5% level. (i.e. the dreaded $p < 0.05$).

2.2 Statistical inference for the difference between observed and expected prevalence

Let $D = p_h - p_s$. Then $D = 0.05$ indicates that the prevalence of the characteristic is five percentage points greater in the host area than in the source area as a whole.

The variance for D is given by:

$$V(D) = V(p_h) - 2Cov(p_h, p_s) + V(p_s)$$

where $V(p_h)$, $Cov(p_h, p_s)$ and $V(p_s)$ are as defined above. A 95% confidence interval for D is given by:

$$D \pm 1.96\sqrt{V(D)}$$

and a significance test can be based on:

$$\chi^2 = \frac{D^2}{V(D)}$$

which can again be compared to tables of the chi-square distribution with one degree of freedom.

3 Statistical methods for comparing host and non-host areas.

When comparing host and non-host areas the ratio and difference measures are $R = p_h/p_{nh}$ and $D = p_h - p_{nh}$. The variance for these measures is given by:

$$V(\ln(R)) = \frac{V(p_h)}{p_h^2} + \frac{V(p_{nh})}{p_{nh}^2}$$

$$V(D) = V(p_h) + V(p_{nh})$$

where

$$V(p_{nh}) = \frac{p_{nh}(1 - p_{nh})}{N_{nh}}$$

and $V(p_h)$ is as defined above.

Confidence intervals and significance tests can be computed using the same formulae given in section 2, but with these modified versions of R, D, $V(\ln(R))$ and $V(D)$.

The methods described here for comparing host and non-host areas are also appropriate for comparing host and "almost-host" areas.

4 Pooling results over sites

The above calculations will be carried out for each site and each characteristic. To summarise these results it would be reasonable to compute an overall summary effect for each characteristic, in effect pooling the information from the various sites. There are several possibilities here but the simplest approach is described below.

4.1 Ratio measures.

Let R_i denote the ratio measure (host v source or host v non-host) for some characteristic for the i^{th} site. Let $W_i = 1/V(ln(R_i)$, the inverse of the variance for the natural logarithm of the ratio for the i^{th} site and let

$$\overline{L} = \frac{\sum_i W_i ln(R_i)}{\sum_i W_i}$$

then $V(\overline{L}) = (\sum_i W_i)^{-1}$ and an overall summary measure of the ratio of host to source area prevalence is given by: $\overline{R} = \exp(\overline{L})$ and a 95% confidence interval is given by:

$$\exp(\overline{L} \pm 1.96\sqrt{V(\overline{L})}.$$

A one degree of freedom chi-square test can be based on the statistic

$$\chi^2 = \frac{\overline{L}^2}{V(\overline{L})}.$$

4.2 Difference measures

Let D_i denote the difference measure (host v source or host v non-host) for some characteristic for the i^{th} site and now let $W_i = 1/V((D_i)$, denote the inverse of the variance for difference measure for the i^{th} site. An overall summary measure of the difference between host and comparison area prevalence is:

$$\overline{D} = \frac{\sum_i W_i D_i}{\sum_i W_i}$$

and $V(\overline{D}) = (\sum_i W_i)^{-1}$. A 95% confidence interval is given by

$$\overline{D} \pm 1.96\sqrt{V(\overline{D})}.$$

A one degree of freedom chi-square test can be based on the statistic

$$\chi^2 = \frac{\overline{D}^2}{V(\overline{D})}.$$

4.3 Heterogeneity

The ratio and/ or difference measures may vary considerably over site, in which case it may not be sensible to compute a pooled summary effect. Al-

though there are statistical tests available for testing heterogeneity of effects across sites these will probably not be useful here, because, with the large sample sizes likely to be available in this study, even trivial heterogeneity will be declared statistically significant. Some judgement will have to be exercised in order to decide whether pooling is sensible. If the ratio or difference measures do seem to vary markedly over sites then it would be interesting to try and relate this to characteristics of the sites (e.g rural v urban may be one site characteristic). This could be achieved by grouping the sites into sub-groups defined according to certain characteristics, checking whether the effect measures seemed relatively homogeneous within subgroups, computing pooled summary estimates for each subgroup and comparing these pooled subgroup estimates.

Other more formal statistical methods may also be relevant here but this would depend to some extent on the software you have available and the number and type of site characteristics considered relevant. This could be discussed at a later date after you have had a preliminary look at the data.