DEPARTMENT OF ECONOMICS AND FINANCE

COLLEGE OF BUSINESS AND ECONOMICS

UNIVERSITY OF CANTERBURY

CHRISTCHURCH, NEW ZEALAND

The Effect of Neighborhood Diversity On Volunteering: Evidence From New Zealand

Jeremy Clark and Bonggeun Kim

WORKING PAPER

No. 09/2009

Department of Economics and Finance
College of Business and Economics
University of Canterbury
Private Bag 4800, Christchurch
New Zealand

The Effect of Neighborhood Diversity On Volunteering: Evidence From New Zealand[*]

Jeremy Clark[a] Bonggeun Kim[b]

A growing empirical literature has found that neighborhood heterogeneity lowers people's likelihood of contributing to public goods, but has struggled to address the issues raised by definition of neighborhood size or endogenous neighborhood location. We show that the estimated effect of concave neighborhood characteristics like heterogeneity on outcomes of interest may be biased if boundaries are defined too broadly, or underestimated if they are defined too narrowly. We also argue that fixed effects panels that follow neighborhoods over time may address problems of endogenous self-selection between neighborhoods if sorting takes place in the manner of the Tiebout hypothesis. We apply both points using three rounds of New Zealand census data to test whether volunteering rates are lowered by neighborhood heterogeneity by race/ethnicity, birthplace, income or language. We find boundaries matter, and that only ethnic/racial heterogeneity is robustly associated with lower volunteering.

Key words: heterogeneity, neighborhood effects, volunteering
JEL Classification: D13, D64, H31

[a] Author for correspondence. Department of Economics and Finance, University of Canterbury, Private Bag 4800, Christchurch, 8140 New Zealand. Email: jeremy.clark@canterbury.ac.nz Phone: 011 643 364 2308 Fax: 011 643 364 2635

[b] Department of Economics, Seoul National University, 599 Gwanak-no, Gwanak-gu, Seoul 151-746, Korea. Email: bgkim07@snu.ac.kr Fax: 82-2-886-4231

1. Introduction

Do individuals in communities that become more heterogeneous lose concern for the welfare of others? Support for this provocative claim has emerged in the past decade over various dimensions of "heterogeneity" and manifestations of "concern for others." To address this question empirically, researchers have commonly tested for an effect of a concave neighborhood characteristic (heterogeneity), on people's propensity to give time or money to public goods, become members of organizations, return census forms, express trust in others, and so on. More generally, researchers have used a similar approach to test for the effect of various concave neighborhood characteristics on various outcomes of interest. Examples include the effect of neighborhood income inequality on mortality rates (Lynch et al. 1998, Mellor and Milyo 2001, Deaton and Lubotsky 2003, Lochner et al. 2001), self reported health (Blakely et al. 2002, Mellor and Milyo 2002), and homicide rates (Mellor and Milyo 2001), or the effect of neighborhood racial diversity on population and economic growth (Rappaport 1999, Alesina and La Ferrara 2005), and adolescent sexual activity (Brewster 1994).[1]

In this paper we make two contributions. By using unusually fine geographical units and a panel of those units, we address two potential problems in previous studies: neighborhood size and endogenous neighborhood choice. Endogenous neighborhood choice may be characterized by a Tiebout-style hypothesis, where people self-select to live in areas with others who share their preferences regarding the optimal trade-off between private consumption and public goods provision (Tiebout 1956). Thus, if people who are less inclined to volunteer are attracted to locate in urban centres, which tend to be more heterogeneous, then the literature's estimated cross sectional effects of heterogeneity on volunteering may be exaggerated. Similarly, the size of neighborhood used to estimate the effect of a concave characteristic (like heterogeneity) on people's behavior may introduce bias if it is defined too broadly. This is because larger unit analysis is likely to incorrectly capture the effect of differences between constituent smaller units due to the concavity of the heterogeneity measure. In contrast, if the size of neighborhood is defined too narrowly to

[1] There is a larger empirical literature testing for the effect of *non-linear* (but not necessarily concave) neighborhood characteristics on various outcomes, covering a wider range of applications (for surveys, see Deaton (2003) or Durlauf (2004)). While neighborhood boundary choice may affect results for all non-linear measures, we have focused here on the consequence of using concave neighborhood measures in particular.

include the effects of heterogeneity in adjacent areas, the lower bound of heterogeneity's total effect can still be estimated. We address both endogenous location and size effects to further the empirical investigation of the effects of neighborhood heterogeneity on people's contribution of time to public goods, or volunteering. We use a heretofore untapped data source that provides several advantages over preceding studies: New Zealand census data on volunteering rates at the unusually fine levels of "meshblock" (= 100 people) and "area unit" (= 2000 people) for 1996, 2001 and 2006. We test whether heterogeneity of ethnicity/race, languages spoken, birthplace, or household income affects New Zealander's likelihood of volunteering. Questions regarding volunteering were asked of all New Zealanders in 1996, 2001 and 2006, enabling us to construct both pooled cross section and neighborhood fixed effects regressions for the entire country. The New Zealand census releases an unusually comprehensive list of covariates for all three years, allowing our cross section regressions to better control for confounding neighborhood characteristics such as deprivation, crime, housing and employment status, that may be correlated with heterogeneity. By using two levels of neighborhood boundary, we address the sensitivity of our empirical findings to neighborhood size and the robustness of previous studies using broader units. Our fixed effects regressions, following neighborhoods over time, can go part way to addressing residual endogeneity in our cross section analysis, based on a Tiebout–style argument that unobserved attitudes towards volunteering may remain stable within neighborhoods over time even as the people in them come and go.

The rest of the paper will proceed as follows. In Section 2, we review the growing empirical literature investigating the effect of neighborhood diversity on public good contributions and trust. In Section 3, we present a simple model that highlights the importance of neighborhood size when estimating the effects of concave neighborhood characteristics on outcomes of interest. In Section 4 we present descriptive statistics regarding volunteering and various measures of heterogeneity in New Zealand, followed by our estimation methods and results. In Section 5 we provide a discussion and conclusions.

2. Diversity, public goods and trust

In recent years researchers have examined the effects of increased neighborhood heterogeneity by race, ethnicity, education, income or first language, on an individual's propensity to volunteer, contribute to fundraisers, be a member of any organization, trust others or support welfare programs.[2] While the bulk of empirical studies have been carried out using data from the United States, others have used surveys from Australia, Kenya, Sweden, and the United Kingdom (see below). The most common approach has been to regress individuals' survey responses on individual and neighborhood level characteristics, with the latter separately taken from census data for the neighborhood or region in which the respondents live.

A selective summary of this literature might suggest that there is indeed a robust negative relationship between heterogeneity and support for public goods and trust in others. Alesina and La Ferrara (2000, 2002), using pooled cross sectional data from multiple years of the U.S. General Social Survey, find that increased neighborhood heterogeneity of income or race lowers an individual's probability of reporting membership in any organization, or of agreeing that "most people can be trusted." Costa and Kahn (2003a), using pooled cross section data from two years of the U.S. Current Population Survey (CPS), find that increased heterogeneity of income or birthplace lowers an individual's probability of membership in any organization or of volunteering. Costa and Kahn (2003b) using the CPS and the DDB Lifestyle Survey, find that increased racial heterogeneity lowers individuals' probability of volunteering. Vigdor (2004) finds that U.S. census tracts that were more heterogeneous in race, age or educational attainment in 2000 had lower response rates of households mailing in completed census forms. Returning such forms can be seen as a local public good, because local public funding depends on enumerated census tract population. Putnam (2007), using responses from the U.S. Social Capital Community Benchmark Survey of 2000, finds that individuals in more racially heterogeneous census tracts were less likely to give to charity or volunteer, trust others (whether of their own or other races), register to vote, or be optimistic that others would cooperate in dilemmas of collective action. Finally, Luttmer (2001), again using pooled cross section data from multiple years of the General Social Survey, finds that support for government welfare spending is lower in more racially heterogeneous states, and that this effect is significant in explaining some of the variation in generosity of welfare across states.

[2] Political scientists such as Robert Putnam (2007) have emphasized the effects of heterogeneity on "social capital", or peoples' beliefs and actions that contribute to "social networks and the associated norms of reciprocity and trustworthiness."

While the bulk of the adverse findings regarding heterogeneity have come from the United States, a limited number of papers have found similar results elsewhere, particularly related to trust. Leigh (2006), using the 1997/98 Australian Community Survey and 1996 Australian census data, finds that increased neighborhood heterogeneity of country of birth or of language spoken at home lowers the probability of individuals trusting their neighbors. Letki (2008), using data from the British Home Office Citizenship Survey of 2001 and census, finds that increased ward level racial heterogeneity lowers individuals' trust in their neighbors. Gustavsson and Jordahl (2008), using the 1994 and 1998 Swedish Election Studies Panel and county level census data, find that increased income inequality in the lower half of the income distribution, or in the proportion of a respondent's county that is foreign born, lowered reported trust in others. Finally, Miguel and Gugerty (2005), using an NGO-funded survey of schools in rural Kenya, find that local ethnic heterogeneity is associated with sharply lower voluntary school fundraiser contributions, resulting in lower quality primary schools.

Theoretically, the negative effects of heterogeneity on people's trust of others or contributions to public goods has been attributed to their innate preference to interact with others like themselves, which can cause social networks and the trust they generate to atrophy as dissimilarities increase (Alesina and La Ferrara 2000, 2002, Putman 2007). People may be less likely to "internalize" the benefits they bestow on the community at large by contributing to public goods if they perceive less similarity between themselves and that community (Vigdor 2004). Linguistic heterogeneity in particular may increase the costs of communication and reduce of quality of information exchanged in networks, making investments in such networks less attractive (Leigh 2006). Ethnic or cultural heterogeneity may also reduce the ability of communities to impose negative social sanctions for free riding across ethnic lines (Miguel and Gugerty 2005). Some have tried to organize these various causal mechanisms via preferences, strategies, and production (Alesina and La Ferrara (2005), Habyarimana et al. (2007).

While the above (selective) summary might suggest conclusive evidence that heterogeneity corrodes people's trust in others and their contributions towards public goods, a closer inspection of this literature show the results to be less robust, and more problematic than they first appear.

Regarding robustness, papers testing for the effects of different kinds of heterogeneity often find that some kinds matter, but others do not, or that multiple kinds may matter when tested individually, but not when tested jointly. And the type of heterogeneity that affects behavior or trust seems to vary from study to study. For example, Alesina and La Ferrara (2002) find that higher neighborhood racial heterogeneity (white, black, Asian etc) lowered trust in other people, but higher heterogeneity of ethnic origin did not, while higher income inequality lowered trust when racial heterogeneity was excluded, but had no significant effect when it was included. Similarly, Alesina and La Ferrara (2000) find that while income, racial and ethnic heterogeneity all lowered the probability of group membership when entered separately, only income heterogeneity mattered when all three measures were included (Alesina and La Ferrara 2002). And while Letki (2008) finds that higher racial heterogeneity lowers trust in the United Kingdom, it has no effect on people's likelihood of formal or informal volunteering, unlike in the United States (Putnam (2007) and Costa and Kahn (2003b)). Again, while higher *birthplace* heterogeneity lowers trust in others in Sweden or Australia (Gustavsson and Jordahl 2008, Leigh 2006), higher *ethnic* heterogeneity (i.e. birthplace of ancestors) in Sweden does not.

Regarding problems, many of the studies on heterogeneity have been based on cross sectional data (Putnam 2007, Alesina et al. 1999, Letki 2008, Leigh (2006), Vigdor (2004), Miguel and Gugerty 2005). They thus cannot be sure that effects attributed to heterogeneity are not instead caused by omitted variables that are correlated with heterogeneity. Letki (2008) in particular argues that neighborhood deprivation, poverty and crime may correlate with ethnic diversity yet be inadequately captured is many preceding studies, making diversity wrongly appear responsible for social withdrawal. Cross sectional studies also cannot determine whether it is the *level* of heterogeneity or *changes* in the level of heterogeneity that is affecting people's behavior.[3]

There remains a problem, however, which heterogeneity studies have not generally recognized: how is the (often constrained) choice of neighborhood size affecting results? The coarseness of neighborhoods used has varied enormously. Gustavsson and Jordahl (2008) use Swedish counties, which contains 200,000 – 300,000 or even over one million people. Alesina and La Ferrara (2000, 2002), Costa and Kahn (2003a) and Luttmer (2001) use a

[3] Luttmer (2001), Costa and Kahn (2003a, 2003b), and Alesina and La Ferrara (2000, 2002), construct pseudo panels of cross sectional survey data, which rely for legitimacy on the representativeness of each wave of the survey. Poterba (1997) uses panel data at the state level. Of the papers we have identified, only Gustavsson and Jordahl (2008) use true panel data at the individual level, using survey data.

respondent's US Metropolitan or Primary Metropolitan Statistical Area (MSA/PMSA), which contain a urban core of at least 50,000 and surrounding suburbs and affiliated towns. Letki uses the U.K census level of ward, which contain anywhere from hundreds to over 30,000 people. Alesina et al. (1999) uses U.S. county. Leigh (2006) uses the Australian postal area, typically containing 20,000 people. Since heterogeneity of income, race etc. can vary dramatically in just a few city blocks, the heterogeneity people experience most intimately may vary widely between areas of these coarsely defined neighborhoods. At the smaller end, Vigdor (2004) and Putnam (2007) define neighborhood at the U.S. census tract level, which commonly involves three to five thousand people. We turn next to a simple model to illustrate why the boundary of neighborhood used in these studies matters. The issue of endogenous neighborhood choice is addressed subsequently in Section 4.3.

3. Who are the people in your neighborhood? A model

In this section, we present a simple model that illustrates the problems that can arise when researchers use overly broad or narrow neighborhood boundaries when estimating the effect of a concave neighborhood characteristic (e.g. heterogeneity) on an outcome of interest (e.g. volunteering). Consider a society with heterogeneity defined in terms of ethnicity, and assume for simplicity that there are only ethnicities 1 and 2. Assume next that the society can be divided into a number of "small" neighborhoods, each of equal size. We will start by assuming that people's likelihood of volunteering is affected only by the heterogeneity they perceive in their immediate small neighborhood, and relax this assumption later. Each small neighborhood *i* can then be defined as one of *n* constituent parts of a "large" neighborhood *j*. Following the literature already cited, we shall assume that ethnic heterogeneity can be correctly captured using a fragmentation index, though our argument holds for any concave measure.[4] Ethnic fragmentation x_{ij} can be constructed for each small neighborhood *i* within large neighborhood *j*, and expressed as the product of the two ethnicities' shares:

$$x_{ij} = [1-\theta_{1ij}^2 - \theta_{2ij}^2] = [1-\theta_{1ij}^2 - (1-\theta_{1ij})^2] = 2\theta_{ij}(1-\theta_{ij})\,. \quad (1)$$

[4] Other concave measures used in studies of neighborhood effects include the Gini coefficient, Theil entropy index, Atkinson deprivation index, coefficient of variation, or neighborhood entropy index (see Hansmann and Quigley 1982). We have repeated the empirical analysis reported in this paper using entropy in particular, and found very similar results to those that follow.

Here $\theta_{ij} \equiv \theta_{1ij}$ is ethnicity 1's share of the population in small neighborhood *i* of large neighborhood *j*. With just two ethnicities, the fragmentation index reaches its maximum value at $\theta_{ij} = 0.5$. In the same way, we can also construct a fragmentation index x_j for the large neighborhood. This will be the product of the two ethnicity's shares in the large neighborhood, but can be equivalently expressed as each ethnicity's average share over the n constituent small neighborhoods.

$$x_j = [1 - \theta_{1j}^2 - \theta_{2j}^2] = [1 - \theta_{1j}^2 - (1-\theta_{1j})^2] = 2\theta_j(1-\theta_j)\,, \qquad (2)$$

where $\theta_j \equiv \theta_{1j} = \dfrac{\sum_{i=1}^{n}\theta_{1ij}}{n}$ is ethnicity 1's share in large neighborhood *j*. From the (strict) concavity of x_{ij} and x_j in (1) and (2), it follows from Jensen's inequality that the fragmentation of the large neighborhood will be (strictly) greater than the mean fragmentation of the *n* constituent small neighborhoods, or

$$E(x_{ij}) = E(f(\theta_{ij})) < f(E(\theta_{ij})) = f(\theta_j) = x_j. \qquad (3)$$

That is, as long as fragmentation varies across small neighborhoods, heterogeneity will appear greater, the larger the neighborhood over which it is defined. Intuitively, total fragmentation in a large neighborhood comes from heterogeneity *within* each of its small neighborhoods, but also from differences in heterogeneity *between* them. More formally, we define "between" heterogeneity x_j^B as the residual difference between the fragmentation index of large neighborhood *j*, and the average fragmentation index of its *i* = 1,…,*n* constituent small neighborhoods, the latter equivalent to "within" heterogeneity x_j^W :

$$x_j^B = x_j - E(x_{ij}) = x_j - \frac{\sum_i x_{ij}}{n} = x_j - x_j^W\,. \qquad (4)$$

This mathematical discrepancy may have an empirical consequence: a study that uses large neighborhood boundaries may bias the relationship it finds between heterogeneity (or indeed, any concave neighborhood characteristic) and people's behavior.

To see this, consider a benchmark regression model that correctly recognizes that changes in small neighborhood heterogeneity x_{ij} can affect small neighborhood volunteering rates y_{ij} directly, or indirectly via the effects on heterogeneity between small neighborhoods x_j^B:

$$y_{ij} = \alpha + \beta_W x_{ij} + \beta_B x_j^B + u_{ij}\,. \qquad (5)$$

Here β_W and β_B are the population regression coefficients of within and between neighborhood heterogeneity, respectively, and u_{ij} is a pure random error. Compare to this an empirical study that uses only small neighborhood boundaries, and regresses y_{ij} on x_{ij} alone, yielding a population regression coefficient β_S. From comparison with (5), β_S may be thought of as a biased estimate of the within effect of heterogeneity, β_W because of the omission of x_j^B. Using a standard result from omitted variable bias,

$$\beta_S = \beta_W + \beta_B \frac{\operatorname{cov}(x_{ij}, x_j^B)}{\operatorname{var}(x_{ij})}. \tag{6}$$

If heterogeneity between small neighborhoods has no effect on volunteering within small neighborhoods, or $\beta_B = 0$, the total effect of heterogeneity will come from the within effect, and be captured by β_S without bias. More generally, if between heterogeneity does affect volunteering, the total (within and between) effect will still be captured by β_S without bias. To see this, note that a change in y_{ij} caused by a unit change in heterogeneity x_{ij} is the sum of the (direct) within effect, $\beta_W \Delta x_{ij}$, and the (indirect) between effect, $\beta_B(\Delta x_j^B / \Delta x_{ij})\Delta x_{ij}$. The total effect on y_{ij} is thus $(\beta_W + \beta_B(\Delta x_j^B / \Delta x_{ij}))\Delta x_{ij}$ or $(\beta_W + \beta_B(cov(x_{ij}, x_j^B)/var(x_{ij})))\Delta x_{ij}$. From (6), the latter can be re-expressed as $\beta_S \Delta x_{ij}$.

In contrast, empirical studies that use only large neighborhood boundaries, and regress y_j on x_j, will have problems. Returning to our benchmark, a simple aggregation over the correct small neighborhood specification in (5) yields the correct large neighborhood specification:

$$y_j (= \frac{\sum_i y_{ij}}{n}) = \alpha + \beta_W x_j^W + \beta_B x_j^B + u_j \quad . \tag{7}$$

By adding and subtracting a common term, (7) can be re-expressed as:

$$\begin{aligned} y_j &= \alpha + \beta_W (x_j^W + x_j^B) + (\beta_B - \beta_W) x_j^B + u_j \\ &= \alpha + \beta_W (x_j) + (\beta_B - \beta_W) x_j^B + u_j \end{aligned} \tag{8}$$

From comparison with (8), we can see that a study using only large boundaries is missing a term involving x_j^B. The resulting estimated effect of large neighborhood heterogeneity x_j on y_j, β_L, can be expressed as

$$\beta_L = \beta_W + (\beta_B - \beta_W) \frac{\operatorname{cov}(x_j, x_j^B)}{\operatorname{var}(x_j)}. \tag{9}$$

In the special case that between heterogeneity has no economic effect on volunteering ($\beta_B = 0$), β_L will not capture the true effect of within heterogeneity. Instead, if $\text{cov}(x_j, x_j^B) <$ 0 then from (9) it follows that $|\beta_L| > |\beta_W|$, or if $\text{cov}(x_j, x_j^B) > 0$ then $|\beta_L| < |\beta_W|$. The concavity of the heterogeneity measure introduces a sort of measurement error into large boundary regressions. More generally, when $|\beta_B| > 0$, large boundary regressions will continue to provide biased estimates of the within- and total effects of heterogeneity.[5]

Problems of a different sort arise if boundaries are set too narrowly. To see this, we relax the assumption that people's likelihood of volunteering is affected only by the heterogeneity they perceive in their immediate small neighborhood.[6] More formally, rather than assuming that y_{ij} can be affected only directly by Δx_{ij} and indirectly by Δx_j^B, we allow that it can also be affected by $\Delta x_{i'j}, i \neq i'$. In this case the total effect on y_{ij} would be the combined one of $(\beta_W + \beta_B(\Delta x_j^B / \Delta x_{ij}))\Delta x_{ij}$ *plus* the effect of $\Delta x_{i'j}, i \neq i'$. As a result, small boundary regressions will themselves yield downward biased estimates ($(\beta_W + \beta_B(\Delta x_j^B / \Delta x_{ij}))\Delta x_{ij} = \beta_S \Delta x_{ij}$) of heterogeneity's total effect. Unfortunately, bad news for overly-small boundary regressions does not translate into good news for larger boundary regressions. β_L will become favorable in the sense of better capturing this additional effect of within heterogeneity from surrounding small neighborhoods $\Delta x_{i'j}, i \neq i'$ on y_{ij}, but it will still be biased because of the omission of a term involving x_j^B as in equations (8) and (9). For example, in the special case that between heterogeneity has no economic effect on volunteering ($\beta_B = 0$), β_L will remain biased in capturing the within effect of Δx_{ij} for all *i* small neighborhoods on y_{ij}.

---

[5] In the special case that volunteering is identically affected by within and between heterogeneity, or $\beta_W = \beta_B$, equations (5) and (7) produce $\beta_S = \beta_L = \beta_W$ and $\beta_W = \beta_B$. It is only in this case that neighborhood size has no effect on the estimated effect of a concave characteristic.

[6] The possibility of heterogeneity outside a small boundary affecting behavior within it also needs to be addressed when we consider endogenous neighborhood choice. We attempt to control for endogenous selection to small or large neighborhood boundaries in Section 4.3.

Our argument might suggest that researchers should estimate concave neighborhood effects using the smallest possible neighborhood size consistent with fully capturing posited effects on outcomes of interest. But researchers are often constrained as to available neighborhood size, or else unsure *ex ante* as to which boundary is just large enough to capture all posited influences. In an additional attempt to capture the within effect of heterogeneity on volunteering, as well as the between- and possible adjacent neighborhood within effects, we will use two boundary levels simultaneously. We will regress small neighborhood volunteering on small *and* associated large neighborhood heterogeneity measures. With the within effect of small neighborhood heterogeneity on volunteering controlled for, the remaining effect of large neighborhood heterogeneity will be equivalent to the sum of between effects and additional within effects of $\Delta x_{i'j}, i \neq i'$. This "hybrid" regression allows us to capture the within and between effects of heterogeneity on volunteering without the measurement error introduced by the concavity of our fragmentation index. We will also present the results of purely small neighborhood specifications to provide a lower bound on heterogeneity's total effect, and the results of purely large specifications to see if the bias we identify in theory makes much difference in practice. By further adding fixed effect estimation of our specification, we will also address in part the issue of endogenous location choice in small (or large) neighborhoods.

4. Empirical Analysis

4.1 The case of New Zealand

Common to other Western nations, New Zealand has experienced a marked increase in social diversity over the past 25 years. Starting as a British colony in the mid-nineteenth century, New Zealand's population was predominantly of British ancestry, with a significant indigenous Maori population (Phillips, 2008). Immigration from other European and Commonwealth countries increased from the time of the second World War, and from neighboring Pacific Island and South East Asian nations. Changes to the Immigration Act of 1987, and the introduction of an ethnicity-blind points system in 1991 was followed by a substantial further diversification of migrants from China, India, and North African and Middle Eastern countries (Phillips, 2008). For more detail about social diversity and volunteering in New Zealand, we turn to the data.

4.1. Data and descriptive statistics

Our data comes from the New Zealand census rounds of 1996, 2001 and 2006. The New Zealand census collects data on an exhaustive list of individual and household characteristics including volunteering activities, ethnicity/race, languages spoken, birthplace and household income. These data are released by Statistics New Zealand at various levels of neighborhood aggregation, right down to the unusually small size of meshblock (≈ 100 people) and area unit (≈ 2000 people). Constant 2006-defined neighborhood geographic boundaries are used for all three rounds to ensure consistency. Our sample is restricted to all those neighborhoods without missing or censored explanatory variables.[7] Over the three years of the census, our pooled sample is 3,504 area units and 49,600 meshblocks in New Zealand.

A description of the dependent and explanatory variables used is provided in Appendix 1, and corresponding descriptive statistics are provided in Appendix 2. To provide the reader with a description of contemporary New Zealand, Figure 1 illustrates the volunteering rate and ethnic, language, birthplace and nominal household income shares for 1996, 2001 and 2006. These shares are population-weighted mean values, based on those meshblocks providing complete observations for our analysis, or our "common sample."[8] Figure 1 suggests that New Zealand is increasing in diversity along each dimension. Moving from shares to fragmentation measures (the equivalent of one minus the Herfindahl Index of concentration), Table 1 provides key descriptive statistics. The population-weighted average proportion of New Zealanders aged 15 or over who reported volunteering at least once

---

[7] In general, we constructed share variables for each neighborhood so as to ensure they were weakly positive and summed to one. In the case of gender, for example, we constructed "ShareFemale" by dividing the frequency of "Number Female" by ("Number Female"+"Number Male"). This assumes that non respondents had the same gender composition as respondents. See Appendix II for details of each variable's construction.

[8] Corresponding descriptive statistics using all meshblocks providing observations for a given variable (our "maximum sample"), are provided in Appendix 3.

outside the household in the previous four weeks was 18.3% in 1996, and then, using a slightly different definition, 15.6% in 2001 and 14.7% in 2006.[9] During this same period,

{Figure 1 about here.}

{Table 1 about here.}

fragmentation by ethnicity/race, languages usually spoken, birthplace, and nominal household income increased. Regarding ethnicity/race, a fragmentation index could conceivably range between 0 and .8 for five categories. The population-weighted mean fragmentation across all meshblocks rose from .347 in 1996, to .352 in 2001, to .378 in 2006. A similar index for language fragmentation, which could range from 0 to .75 over four categories, rose from 0.246 in 1996 to .254 in 2001 to .275 in 2006. The index for birthplace fragmentation, which could range from 0 to .5 over two categories (inside or outside of New Zealand), rose from .293 in 1996 to .300 in 2001 to .329 in 2006. Finally, the index for nominal household income fragmentation, which could range from 0 to .833 over six unadjusted nominal income bands, rose from .746, to .757, to .766.[10]

Another way to illustrate potential correlation between volunteering and fragmentation is to graph 95% confidence intervals of the best fit polynomial relationship between each

[9] For 1996 volunteering was defined as having "Attended Committee Meeting etc Unpaid for Group, Church or Marae." For 2001 and 2006 the definition was changed to be defined as any "Other Helping or Voluntary Work For or Through any Organisation, Group or Marae." For all three years our definition excludes those caring for a child or someone who was ill, elderly, or disabled outside the household. See Appendix Table 1I for more detail. Because of the change in volunteering question, we have repeated all the analysis to follow using only 2001 and 2006 data. The results concerning heterogeneity's effects are very similar to what we report here, with the exception that the evidence for income heterogeneity's (negative) effect is slightly greater, becoming significant even in fixed effects analysis.

[10] Because the six household income bands were not adjusted for inflation between each census, we can only measure how the dispersion of unadjusted nominal incomes across bands has changed over time, not real household incomes.

meshblock's volunteering rate and each dimension of heterogeneity. This is done in Figure 2 using the pooled sample. A clear negative relationship appears between volunteering and ethnic, language and birthplace fragmentation, while a more diffuse inverted U relationship appears for nominal income fragmentation. The non-monotonic relationship between income fragmentation and volunteering might reflect that there exists a degree of inequality beyond which volunteering is depressed, or simply that a correlation does not exist.

{Figure 2 about here.}

Thus, consistent with the findings in the literature, the fall in volunteering rates in New Zealand coincided with increasing heterogeneity by several dimensions. Nonetheless, many other changes were taking place in New Zealand over these years which could have influenced people's decision to volunteer (via their tastes or opportunity costs), or organizations' decisions to demand volunteers (via volunteers' non-wage costs and productivity (Handy and Srinivasan (2005))). We construct measures for many of these confounding factors, which are described in Appendices 1 and 2, including real median household income, and the ethnicity, language, and birthplace shares that underlie our fragmentation measures. Among these variables, the average real median household income across meshblocks rose from NZ$ 37,800 in 1996, to $39,000 in 2001, to $45,000 in 2006. The mean share of females remained steady at 51%, while the mean percentage whose highest education was a bachelor's or honour's degree rose from 8% to 10% to 12%. At the same time, the mean percentage of those aged 15 or over not in the labor force fell from 34% to 33% to 31%. The mean percentage claiming Christian religious affiliation also fell from 67% to 62% to 56%, while the mean percentage claiming no religious affiliation rose from 28% to 31% to 36%. We will try to untangle the effects of these various changes on volunteering rates in the regression analysis that follows.

Finally, returning to the issue of neighborhood boundary, Table 2 compares the mean and standard deviation of meshblock and area unit measures of fragmentation.[11] As predicted in Section 3, the means of all four types of heterogeneity appear greater over area units than over meshblocks. In addition, the standard deviation of neighborhood heterogeneity is consistently lower at the area unit level than at the meshblock level for every measure. This

[11] The descriptive statistics in Tables 1 and 2 use the common sample used for subsequent regression analysis.

suggests that our choice of neighborhood size may indeed affect the empirical relationship we estimate between social heterogeneity and volunteering.

4.3. Estimation strategy and results

In this section, we lay out and implement our strategy for estimating the cross-sectional and longitudinal empirical relationship between social heterogeneity and volunteering. Because we have a wide, shallow panel of many neighborhoods over just three census years, our pooled data contains substantial variation *between* neighborhoods at any point in time,

{Table 2 about here.}

but less variation *within* neighborhoods over time. In line with the vast majority of studies in our literature review, we shall begin by using pooled cross section OLS as our baseline specification. To address the problem of omitted variable bias that attends cross section analysis, and to test the robustness of our results, we shall then add two steps. First, we will repeat the baseline cross sectional analysis using various additional groups of control variables. Second, we will switch to fixed effects analysis to address omitted variable bias resulting from location choice under a Tiebout-style hypothesis. Throughout this process, we shall provide (small) meshblock regressions that provide either an unbiased or lower-bound estimate of heterogeneity's total effect on volunteering. We shall also provide larger area unit regressions for comparison, and finally hybrid regressions. Hybrid regression coefficients may be thought of as decomposing the total effects of heterogeneity into within and between effects if volunteering is affected only by heterogeneity in x_{ij}. If volunteering is also affected by heterogeneity in $x_{i'j}, i \neq i'$, hybrid regressions may be thought of as providing an upper bound of heterogeneity's total effect, (assuming area units contain all relevant heterogeneity), given by the sum of within and between coefficients. In such cases, the lower bound of total effects is given by the single heterogeneity coefficient in the meshblock regressions.

Beginning with our baseline cross sectional analysis, we run regressions of the form

$$y_{ijt} = X'_{ijt}\beta + u_{ijt} \tag{10}$$

where y_{ijt} is meshblock i's volunteering rate in area unit *j* in year *t*. X_{ijt} is a vector of neighborhood characteristics, year dummies, and social heterogeneity measures, while u_{ijt} is a random error. In each case, we regress volunteering rates on one type of heterogeneity at a time, along with its underlying level variables (e.g. ethnic shares for ethnic fragmentation, language shares for language fragmentation, median household income for income fragmentation, etc.). To these we add the baseline covariates of share female, median age, population density, mean household size, share married, shares of families comprised of couples with children, and couples without, and year dummies.[12]

Table 3 provides the results. Column (1) shows our baseline estimate of meshblock ethnic heterogeneity's effect on meshblock volunteering rates with controls for (meshblock) ethnic affiliation shares and baseline covariates. The estimated coefficient on ethnic fragmentation (-.128) implies a relatively strong negative effect of this type of heterogeneity on volunteering. In particular, a 10 percentage point increase in meshblock ethnic fragmentation is estimated to decrease the (meshblock) volunteering rate by 1.3 percentage

{Table 3 about here.}

points. This effect is only moderately overstated when area units are used in place of meshblocks in column (2), where a 10 percent increase in area unit ethnic/racial fragmentation decreases the (area unit) volunteering rate by 1.4 percentage points. Finally, from our hybrid specification in column (3), a 10 percentage point increase in heterogeneity within meshblocks decreases (meshblock) volunteering by 1.1 percentage points. At the same time, a 10 percent increase in corresponding area unit heterogeneity decreases meshblock volunteering by .4 percentage points. This indicates the existence of between effects and/or the effect of heterogeneity in other meshblocks $\Delta x_{i'j}, i \neq i'$.[13]

---

[12] We have also added ethnic share composition to the baseline covariates when examining the effects of birthplace or income fragmentation on volunteering. This is because of the clear effect that Maori ethnic affiliation has on volunteering rates. Ethnic shares remain omitted when examining the effects of language fragmentation, because ethnic and language shares are highly correlated.

[13] The 1.05 ($\approx$ 1.1) within, and .40 between effects of heterogeneity can be related to the 1.28 total effect as follows. If we assume heterogeneity in other meshblocks has no effect, then on average a ten percentage point increase in ethnic fragmentation within meshblocks raises ethnic fragmentation between meshblocks in the associated area unit by 5.75 percentage

The estimated effects of language, birthplace and household income heterogeneity are similarly provided in columns (4) – (12) of Table 3. Language and birthplace heterogeneity are again strongly negatively associated with volunteering rates under the baseline specification. From column (4), language fragmentation's (lower bound of) total effect on volunteering is a 2.8 percentage point drop, though it appears as a 3.7 percentage point drop when larger area units are used. Similarly, from column (7), birthplace fragmentation's (lower bound of) total effect on volunteering is a 1.1 percentage point drop, which appears as a .3 percentage point drop using area unit neighborhoods. Finally, nominal household income band fragmentation's (lower bound) total effect on volunteering in column (10) is a slight .2 percentage point drop, which appears as a 1.1 percentage point drop using area units. Thus, our baseline cross section results might suggest that New Zealand's shifting immigration and tax policy has been responsible for a drop in New Zealander's tendency to contribute time towards public goods. These initial results also suggest that using broader neighborhood boundaries can either inflate or attenuate heterogeneity's estimated effects by as much as five times.

While our meshblock estimates of the effects of each type of heterogeneity on volunteering are almost uniformly negative, this could simply reflect the omission of other influences on volunteering that are correlated with heterogeneity. Omitted factors could include variation in religious affiliation, neighborhood deprivation, labor force status, or education. We have also yet to test whether one type of heterogeneity affects volunteering when other dimensions of heterogeneity (and their underlying share variables) are controlled for. Thus, in Table 4 we extend our cross sectional analysis to include groups of other confounding variables one at a time.[14] These groups are: 1) religious affiliation: Christian, other, and no religion affiliation rates, 2) neighborhood deprivation: home ownership rates, median number of bedrooms, crime rates, and percentage of individuals receiving single parent domestic benefits, 3) employment status: shares in full time work, part time work,

---

points, or $\Delta x_j^B / \Delta x_{ij} = cov(x_{ij}, x_j^B) / var(x_{ij})$ = .575. The total effect is then 1.05 + .40(.575) – 1.28. If we assume heterogeneity in other meshblocks is relevant, 1.28 becomes the lower bound of total effects, including the effect of $\Delta x_{i'j}, i \neq i'$.

[14] High degrees of correlation between various covariates precluded us from including all clusters simultaneously.

unemployed, and not in the labor force, 4) education levels: the share of individuals lacking minimum high school qualifications and the share with bachelor's or (additional year) honour's degrees, and 5) including all heterogeneity measures simultaneously, together with their underlying share variables. Note that care must be taken in evaluating the effects of each type of heterogeneity when all are included simultaneously in 5), because they (and their underlying share variables) may be highly correlated.

While many of our baseline and added covariates explain variation in volunteering rates,[15] we focus in Table 4 on showing the direct (remaining) effect of each type of heterogeneity on volunteering. Column (1) of Table 4 shows the (lower bound of) total effect of meshblock ethnic heterogeneity on meshblock volunteering as each group of confounding variables is added to the baseline covariates of Table 3. In each case, ethnic/racial heterogeneity retains a significant, negative effect on volunteering of roughly 1 percentage point. Language heterogeneity similarly retains a robust and relatively strong negative effect as other covariates are included. From column (4), a 10 percentage point increase in

{Table 4 about here.}

meshblock language fragmentation results in a roughly 2 to 2.5 percentage point drop in the volunteering rate.

Birthplace heterogeneity (column 7) similarly retains a negative effect as covariates are added, similar in magnitude to ethnic heterogeneity (1 percentage point), with the exception of the case where all types of heterogeneity and underlying share variables are entered simultaneously. There a 10 percentage point increase in meshblock birthplace fragmentation *raises* volunteering by .2 percentage points. Finally, nominal household income band

---

[15] Those covariates consistently positively related to volunteering rates were share with Maori ethnic affiliation, Maori and Samoan language shares, median age, share married, share of families that had couple with kids, and less so families that had a couple without kids, share with Christian or other religious affiliation, share who owned own home, median number of bedrooms, share with bachelors or honours degrees, and share employed part time. Those covariates consistently negatively related to volunteering rates were share with Asian or MELAA ethnic affiliation, English or "other" language share, household size, population density, share with no religious affiliation, share of families that were single parent, and share employed full time.

heterogeneity is similar to birthplace heterogeneity in retaining a negative effect on volunteering as covariates are added, except when all types of heterogeneity are included. However, the magnitude of the effect of income fragmentation on volunteering is smaller than for other types of heterogeneity. It ranges between lowering volunteering by .4 percentage points, to raising it by .1 percentage point when all heterogeneity is considered simultaneously. While it is interesting that birthplace and income heterogeneity no longer lower volunteering rates when all types of heterogeneity are controlled for, the correlation that exists between various types of fragmentation mean that these findings must be treated with caution.

Taken together, the baseline and extended cross sectional evidence so far points strongly to a negative effect of ethnic/racial and language heterogeneity on volunteering, and possibly to a negative effect of birthplace and income heterogeneity as well. Nevertheless, as with any cross section analysis, there may remain unobserved characteristics that are correlated with heterogeneity that are skewing its estimated effects. Even with the inclusion of extensive covariates, omitted variable bias remains a strong possibility because of the Tiebout hypothesis that people will self-select to live in an area with others who share their preferences regarding the optimal trade-off between private consumption and public goods provision (Tiebout 1956). Specifically, if people who are less inclined to volunteer are attracted to live in urban centres, which tend to be more heterogeneous, then the cross sectional effects of heterogeneity on volunteering will be exaggerated. We therefore move to fixed effects analysis in an attempt to better control for unobserved characteristics like people's attitude towards volunteering or towards heterogeneity.

One problem with using fixed effects analysis here is that we are following *neighborhoods* rather than *individuals* over time, and the latter are free to change where they live. Is there any reason to expect that unobserved characteristics like attitude to volunteering would remain constant over time for given neighborhoods, even as the individuals in them come and go? Perhaps ironically, our defence of this proposition comes from the same Tiebout hypothesis that raises the problem of endogenous neighborhood choice in the first place. For with freedom of movement, individuals who come to differ with the local prevailing preferences between private and public good provision may leave, and those who share those preferences may enter. If the effects of heterogeneity found with fixed effects do not correspond with those found in cross section, this might suggest that the cross section effects are spuriously caused by omitted variable bias.

To proceed, we estimate the following volunteering equation using panel data on the meshblocks of New Zealand:

$$y_{ijt} = X'_{ijt}\beta + \alpha_{ij} + \varepsilon_{ijt}. \tag{11}$$

y_{ijt} is the volunteering rate in meshblock *i* within area unit *j* in year *t*, while X_{ijt} contains our set of heterogeneity measures and other control variables previously defined. The α_{ij} are unobservable meshblock-specific fixed effects which may be correlated with X_{ijt}, while ε_{ijt} is a pure random error term. To control for the potential correlation between α_{ij} and X_{ijt}, we apply OLS to the mean-differenced equation

$$y_{ijt} - \overline{y_{ij}} = (X_{ijt} - \overline{X_{ij}})'\beta + \varepsilon_{ijt} - \overline{\varepsilon_{ij}}. \tag{12}$$

Here, for any variable Z, $\overline{Z_{ij}} = \sum_t Z_{ijt} / 3$.[16]

Table 5 presents the results. The control variables included are identical to those used in baseline cross section (Table 3), and the effects are again presented using meshblock, area unit, and hybrid specifications. In general, evidence of an effect of heterogeneity on volunteering has weakened. As shown in column (1), ethnic fragmentation retains its negative effect on volunteering, but the magnitude of the (lower bound of) total effect has fallen from 1.1 percentage points in cross section to .5 percentage points under fixed effects. This suggests that neighborhoods that experience an increase in ethnic heterogeneity also experience a decrease in volunteering rates on average, but the effect is modest after controlling for people's choice of neighborhood location in the manner of the Tiebout hypothesis. In contrast, language fragmentation has lost any significant effect (column (4)). Birthplace fragmentation in

{Table 5 about here.}

column (7) retains a negative but reduced effect, with a (lower bound of) total effect falling from 1.1 to .3 percentage points. Finally, nominal household income fragmentation in

[16] We consider also the more general case that endogenous location choice may occur even for large neighborhoods by adding an α_j term in equation (11), or an unobservable area unit-specific fixed effect which may be correlated with X_{jt}. The fixed effects of our hybrid specification will then control for both α_{ij} and α_j.

column (10) is similar to language fragmentation in losing a significant effect.[17] All of these findings are qualitatively summarized in Table 6.

Methodologically, it is interesting to note that the choice of a larger neighborhood boundary would have substantially affected our (baseline) fixed effects results, just as it did our (baseline) cross section analysis. If we use area units rather than meshblocks, ethnic fragmentation would inflate to a -1.1 rather than -.5 percentage point effect. Language fragmentation would inflate to a -2.1 percentage point effect from no effect. Birthplace fragmentation would attenuate to have no effect, rather than a -.3 percentage point effect. Only household income fragmentation would have a similar result – no effect – under either neighborhood boundary.

5. Discussion and conclusions

This paper has attempted to make two contributions to the growing empirical literature in which researchers estimate the effects of concave neighborhood characteristics like heterogeneity on outcomes of interest, such as people's likelihood of volunteering. With access to unusually fine geographical units, and a panel of those units, we have addressed the problems posed by 1) the researcher's choice of neighborhood boundary, and 2) the researchees' endogenous choice of neighborhood. Regarding boundaries, we show that for concave neighborhood characteristics like heterogeneity, using the smallest possible boundary that includes all posited effects will provide an unbiased estimate of total effect. Using overly-large boundaries will incorrectly capture the effect of differences in heterogeneity between constituent small neighborhoods on the outcome of

{Table 6 about here.}

interest. This can bias estimates (up or down) of heterogeneity's total effect even when such "between heterogeneity" has zero economic effect. Using overly-small boundaries will avoid this bias, but provide only a lower-bound estimate of heterogeneity's total effect by ignoring

[17] This is the single case where restriction of the sample to the two years with an identical volunteering definition, 2001 and 2006, resulted in a different finding. Here in fixed effects analysis, income fragmentation retains a negative effect of .2 percentage points, significant at the one percent level.

the influence of heterogeneity in adjacent neighborhoods. Hybrid approaches that use small and large boundaries simultaneously can also be used to distinguish the effect of heterogeneity within a small neighborhood, vs. the combined effects of heterogeneity in other small neighborhoods and of differences in heterogeneity between them.

Regarding endogenous neighborhood choice, we use a Tiebout-style argument to identify a likely problem: people who are less inclined to volunteer may be attracted to locate in (often more heterogeneous) urban centres with others who make similar tradeoffs between private consumption and volunteering. The existing literature's estimated cross sectional effects of heterogeneity on volunteering may thus be exaggerated. We next use the Tiebout hypothesis to identify a potential solution: while individuals may freely choose their small (or large) neighborhoods, the typical attitude towards volunteering in those neighborhoods may remain stable over time as like-minded people enter and contrary-minded people leave. Fixed effects analysis that follows neighborhoods over time may thus be able to control for unobserved but stable characteristics such as attitudes towards volunteering or heterogeneity.

We address both issues in practice by using a heretofore untapped source, New Zealand census data on volunteering over three rounds. We test for a negative relationship between volunteering and heterogeneity by ethnicity/race, language, birthplace, and household income. The New Zealand census is unusual in releasing data in neighborhoods as small as meshblocks ($\approx$ 100 people), and in asking New Zealanders whether they have volunteered in the four weeks prior to the census in 1996, 2001, and 2006. It thus provides an unusual opportunity to carefully examine the effects of neighborhood size and endogenous sorting on estimates of the effect of heterogeneity on volunteering rates.

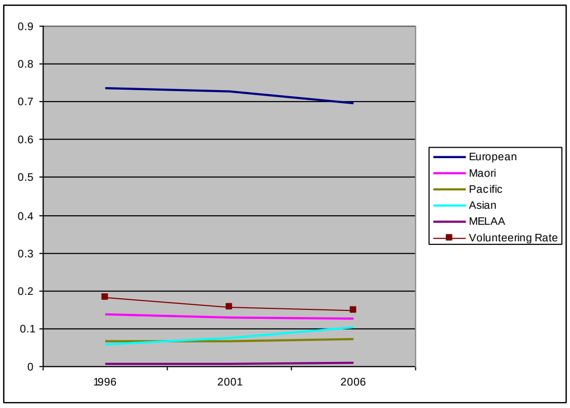
Our baseline cross sectional analysis suggests that diversity indeed discourages volunteering. On average, a ten percentage point increase in ethnic fragmentation lowers the meshblock volunteering rate by at least 1.3 percentage points. The corresponding lower bounds of negative effects for language, birthplace and household nominal income fragmentation are 2.8, 1.1 and .2 percentage points, respectively. Estimated effects tend to be larger when larger area unit boundaries are used (1.4, 3.7, .3 and 1.1 percentage points, respectively). This may be due to a combination of bias and correctly capturing heterogeneity influences external to the meshblock.

Recognizing that our baseline cross section regressions may be suffering from omitted variable bias, we repeat the analysis adding additional groups of covariates relating to religious affiliation, deprivation, employment, and education, as well as including the four types of heterogeneity simultaneously. Extending our cross section covariates with one type of heterogeneity at a time does not greatly affect the magnitude of our negative findings. On the other hand, including all types of heterogeneity and their underlying share variables simultaneously results in ethnic and language fragmentation retaining a negative effect, but birthplace and income fragmentation switching to a significant *positive* effect on volunteering.

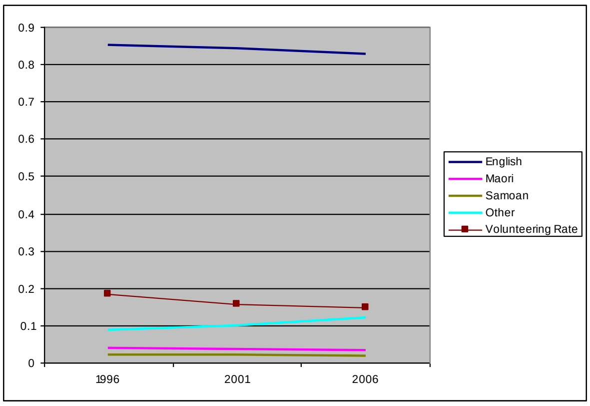
Finally, since even expanded cross sectional analysis still omits unobserved variables that may be correlated to heterogeneity, we repeat our baseline cross sectional analysis using fixed effects that follow meshblocks (or larger area units) over time. Here we find that the evidence of heterogeneity having a negative effect weakens. Compared to analogous cross section regressions, the lower bound total effect of ethnic fragmentation, while still negative, falls from 1.3 to .5 percentage points. The lower bound negative effect of language fragmentation disappears (from 2.8 to 0 percentage points). The lower bound negative effect of birthplace fragmentation remains, but falls from 1.1 to .3 percentage points, while the lower bound negative effect of income fragmentation also disappears (from .2 to 0 percentage points). Interestingly, a coarser level of neighborhood fixed effects analysis would suggest that ethnic and language fragmentation maintain the larger negative effects found in cross section (1.1 and 2.1 percentage points, respectively), but that birthplace fragmentation loses its effect.

Summarizing our findings, the lower bound of total effect of ethnic fragmentation on volunteering is found to be negative in baseline and extended cross section analysis, and weaker but still negative in fixed effects. The lower bound of total effect of language fragmentation is strong and negative in baseline and extended cross section, but non-existent in fixed effects. The lower bound of total effect of birthplace fragmentation is negative in baseline and extended cross section specifications, save where other types of heterogeneity are controlled, and weaker but still significantly negative in fixed effects. Finally, the lower bound of total effect of nominal household income fragmentation is unstable, being relatively minor but negative in baseline and most extended cross section specifications, positive when other types of heterogeneity are controlled for, and non-existent in fixed effects.

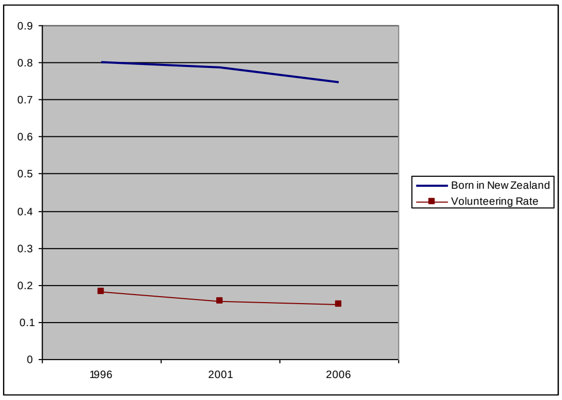
To the extent that fixed effects analysis does not find negative effects from language or income heterogeneity when cross section analysis does, it is unclear whether this is due to fixed effects' better ability to control for unobserved variables, or to its being subject to an errors-in-variable bias that lowers estimated effects towards zero because neighborhood self-selection is not adequately controlled. We are left with fairly conclusive evidence that heterogeneity by race/ethnicity depresses volunteering rates in New Zealand, suggestive evidence that heterogeneity by birthplace may do the same, and inconclusive evidence regarding heterogeneity by language or nominal income.



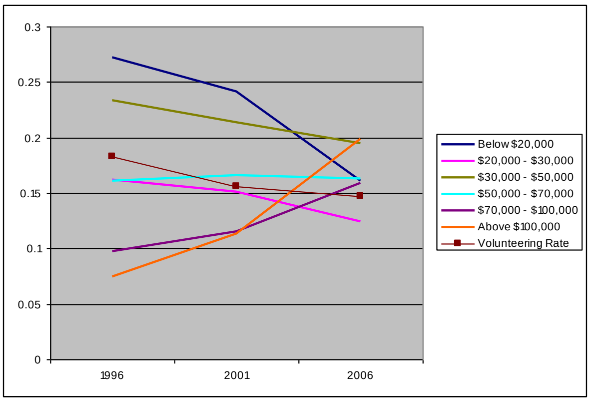
Ethnic Affiliation Shares
(a)



Languages Spoken Shares
(b)

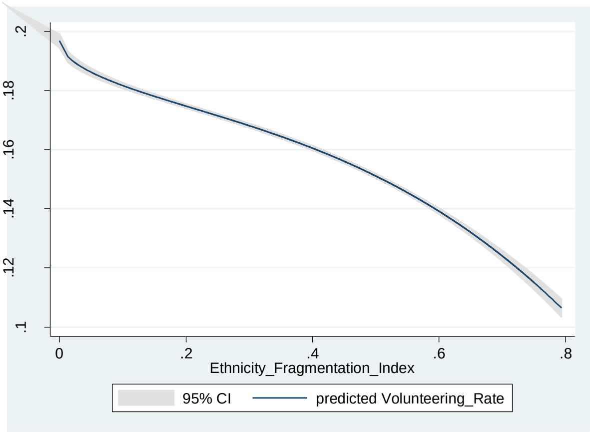


Born in New Zealand Share
(c)

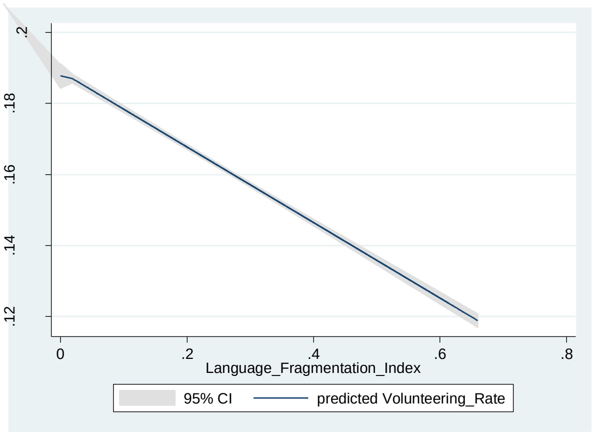


Nominal HH Income Shares (NZ$)
(d)

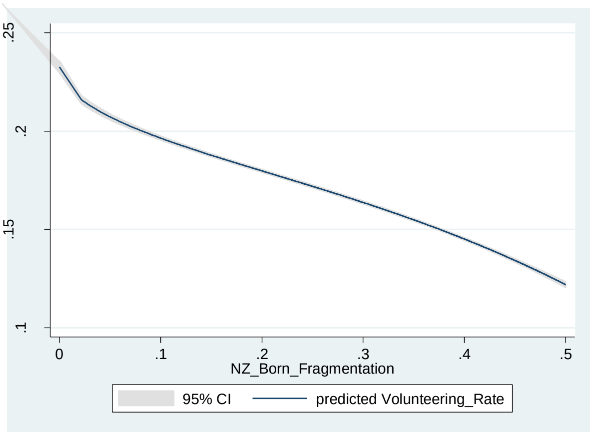
Fig. 1: Population Weighted Mean Meshblock Ethnicity, Language, Birthplace and Nominal Household Income Shares in New Zealand (N_{1996} = 11658, N_{2001} = 18574, N_{2006} = 19368)



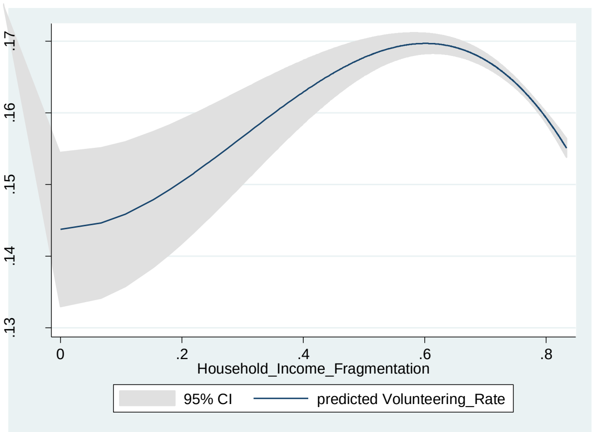
(a)



(b)



(c)



(d)

Fig. 2: Best Fit Polynomial Relationship Between Meshblock Volunteering Rate and (a) Ethnic, (b) Language, (c) Birthplace and (d) Household Income Fragmentation.
Pooled Sample (N = 49,600)

Table 1:
Population weighted means and standard deviations of key variables at the meshblock level, using a common sample.

Census Year	*Variable*				
	Volunteering Rate Mean (St. Dev	Ethnic/Racial Fragmentation Mean (St. Dev)	Language Fragmentation Mean (St. Dev)	Birthplace Fragmentation Mean (St. Dev)	Household Income Fragmentation Mean (St. Dev)
1996 N=11658	.183 (.064)	.347 (.185)	.246 (.136)	.293 (.126)	.747 (.072)
2001 N=18574	.156 (.064)	.352 (.191)	.255 (.141)	.300 (.132)	.757 (.067)
2006 N=19368	.147 (.060)	.378 (.192)	.275 (.143)	.330 (.129)	.767 (.065)

Table 2:
Pooled population weighted means and standard deviations of heterogeneity variables at the meshblock and area unit levels, using a common sample.

Neighbor-hood Size	*Variable*			
	Ethnic/Racial Fragmentation Mean (St. Dev)	Language Fragmentation Mean (St. Dev)	Birthplace Fragmentation Mean (St. Dev)	Household Income Fragmentation Mean (St. Dev)
Small (Meshblock) N=49600	.361 (.190)	.261 (.141)	.310 (.130)	.758 (.068)
Large (Area Unit) N=3504	.402 (.176)	.283 (.127)	.318 (.120)	.797 (.035)

Table 3:
Determinants of volunteering rates: baseline cross section regression (pooled OLS, meshblock N = 49,600, area unit N = 3504).

	(1)	(2)	(3)	(4)	(5)	(6)
Variable	Ethnic Fragment Meshblock	Ethnic Fragment Area Unit	Ethnic Fragment Hybrid	Language Fragment Meshblock	Language Fragment Area Unit	Language Fragment Hybrid
Intercept	.219	.010	.219	.345	.282	.372
	(.018)***	(.019)	(.018)***	(.020)***	(.013)***	(.022)***
Fragmentation: Meshblock	-.128		-.105	-.275		-.244
	(.004)***		(.004)***	(.015)***		(.015)***
Fragmentation: Area Unit		-.135	-.040		-.372	-.068
		(.002)***	(.003)***		(.009)***	(.004)***
Asian Ethnic Share	-.059	.076	-.048			
	(.017)***	(.018)***	(.016)***			
Pacific Ethnic Share	-.051	.070	-.044			
	(.016)***	(.017)***	(.016)**			
Maori Ethnic Share	.157	.317	.161			
	(.016)***	(.017)***	(.016)***			
European Ethnic Share	-.084	.010	-.082			
	(.017)***	(.018)	(.017)***			
English Language Share				-.232	-.350	-.253
				(.020)***	(.011)***	(.020)***
Maori Language Share				.574	.792	.552
				(.009)***	(.006)***	(.009)***
Samoan Language Share				.074	.232	.076
				(.008)***	(.005)***	(.008)***
Female Share	.007	-.007	.010	.016	.089	.019
	(.007)	(.010)	(.007)	(.007)	(.010)***	(.007)
Median Age	.000	.001	.000	.000	.001	.000
	(.000)***	(.000)***	(.000)***	(.000)***	(.000)***	(.000)***
Household Size	-.007	-.009	-.007	-.004	-.007	-.003
	(.001)***	(.000)***	(.001)***	(.001)***	(.000)***	(.001)***
Marriage Share	.102	.024	.099	.106	.031	.101
	(.003)***	(.003)***	(.003)***	(.003)***	(.003)***	(.003)***
Share of Families that are "Couple with Kids"	.037	.171	.037	.039	.221	.041
	(.004)***	(.005)***	(.004)***	(.004)***	(.005)***	(.004)***
Share of Families "Couple with No Kids"	.017	.116	.018	.027	.168	.025
	(.003)***	(.004)***	(.003)***	(.003)***	(.003)***	(.003)***
Population Density	.000	.000	.000	.000	.000	.000
	(.000)***	(.000)***	(.000)***	(.000)***	(.000)***	(.000)***
Year 2001	-.026	-.000	-.026	-.025	.000	-.027
	(.001)***	(.000)	(.001)***	(.001)***	(.000)	(.001)***
Year 2006	-.030	-.000	-.031	-.029	-.000	-.031
	(.001)***	(.000)	(.001)***	(.001)***	(.000)	(.001)***
R^2	.266	.593	.269	.262	.612	.263

Note: ***,**,* represent the levels of statistical significance of 1%, 5%, and 10% respectively. Run on Stata 9.2. Robust standard errors in parentheses.

Table 3 (Cont'd):
Determinants of volunteering rates: baseline cross section regression (pooled OLS, meshblock N = 49,600, area unit N = 3504).

	(7)	(8)	(9)	(10)	(11)	(12)
Variable	Birthplace Fragment Meshblock	Birthplace Fragment Area Unit	Birthplace Fragment Hybrid	H Income Fragment Meshblock	H Income Fragment Area Unit	H Income Fragment Hybrid
Intercept	.080 (.017)***	-.380 (.018)***	.092 (.017)***	-.020 (.017)**	-.374 (.018)***	.028 (.018)
Fragmentation: Meshblock	-.108 (.006)***		-.052 (.006)***	-.019 (.004)***		-.003 (.005)
Fragmentation: Area Unit		-.027 (.00)***	-.087 (.005)***		-.108 (.004)***	-.077 (.010)***
Born in NZ Share	-.031 (.007)***	.073 (.004)***	-.031 (.007)***			
Real Median HH Income				-.000 (.000)***	-.000 (.000)***	-.000 (.000)***
Asian Ethnic Share	-.013 (.016)	.238 (.018)***	-.004 (.016)	.024 (.017)	.236 (.018)***	.023 (.016)
Pacific Ethnic Share	.039 (.016)**	.268 (.017)***	.041 (.016)**	.073 (.016)***	.309 (.017)***	.073 (.016)***
Maori Ethnic Share	.181 (.017)***	.428 (.017)***	.169 (.016)***	.266 (.016)***	.523 (.017)***	.262 (.016)***
European Ethnic Share	.051 (.016)***	.227 (.017)***	.047 (.016)***	.123 (.016)***	.306 (.017)***	.121 (.016)***
Female Share	.016 (.007)**	-.034 (.011)***	.018 (.007)	-.005 (.007)	.049 (.011)***	-.004 (.007)
Median Age	.001 (.000)***	.001 (.000)***	.001 (.000)***	.000 (.000)***	.001 (.000)***	.001 (.000)***
Household Size	-.003 (.001)***	-.005 (.000)***	-.002 (.001)***	-.001 (.001)*	-.003 (.000)***	-.001 (.001)
Marriage Share	.106 (.003)***	.025 (.004)***	.100 (.003)***	.118 (.003)***	.017 (.004)***	.119 (.003)***
Share of Families that are "Couple with Kids"	.043 (.004)***	.233 (.006)***	.047 (.004)***	.049 (.004)***	.377 (.007)***	.050 (.004)***
Share of Families "Couple with No Kids"	.028 (.004)***	.180 (.004)***	.028 (.003)***	.031 (.004)***	.260 (.004)***	.032 (.004)***
Population Density	.000 (.000)***	.000 (.000)***	.000 (.000)***	.000 (.000)***	.000 (.000)***	.000 (.000)***
Year 2001	-.025 (.001)***	.001 (.000)**	-.025 (.001)***	-.024 (.001)***	.001 (.000)***	-.023 (.001)***
Year 2006	-.029 (.001)***	.001 (.000)***	-.028 (.001)***	-.026 (.001)***	.002 (.000)***	-.024 (.001)***
R^2	.251	.560	.256	.242	.580	.243

Note: ***, **, * represent the levels of statistical significance of 1%, 5%, and 10% respectively. Run on Stata 9.2. Robust standard errors in parentheses.

Table 4:
Adding covariates as a robustness check on the effect of heterogeneity on volunteering rates (pooled OLS, meshblock N = 49,600, area unit N = 3504)

Specification	(1)	(2)	(3)	(4)	(5)	(6)
	Ethnic Fragment Meshblock	Ethnic Fragment Area Unit	Ethnic Fragment Hybrid	Language Fragment Meshblock	Language Fragment Area Unit	Language Fragment Hybrid
Basic (Same as Table 3) Meshblock	-.127 (.004)***		-.102 (.004)***	-.267 (.015)***		-.234 (.015)***
Basic (Same as Table 3) Area Unit		-.136 (.002)***	-.041 (.003)***		-.380 (.009)***	-.074 (.004)***
Basic + Religious Affiliation (ChrSh, OthSh) Meshblock	-.109 (.004)***		-.084 (.004)***	-.240 (.015)***		-.201 (.016)***
Basic + Religious Affiliation (ChrSh, OthSh) Area Unit		-.108 (.002)***	-.041 (.003)***		-.307 (.008)***	-.072 (.004)***
Basic + Deprivation (OwnHmSh, MedBedrm, Crime)[a] Meshblock	-.126 (.004)***		-.102 (.005)***	-.256 (.017)***		-.231 (.017)***
Basic + Deprivation (OwnHmSh, MedBedrm, Crime)[a] Area Unit		-.137 (.002)***	-.041 (.003)***		-.369 (.010)***	-.065 (.005)***
Basic + Employment (UnempSh, EmpFTSh, NotLFSh) Meshblock	-.119 (.004)***		-.100 (.004)***	-.265 (.015)***		-.241 (.016)***
Basic + Employment (UnempSh, EmpFTSh, NotLFSh) Area Unit		-.088 (.003)***	-.032 (.003)***		-.237 (.010)***	-.054 (.004)***
Basic + Education Levels (NoQualSh, BHSh) Meshblock	-.124 (.004)***		-.098 (.004)***	-.355 (.015)***		-.322 (.015)***
Basic + Education Levels (NoQualSh, BHSh) Area Unit		-.130 (.002)***	-.043 (.003)***		-.428 (.009)***	-.081 (.004)***
Basic + All Fragment + Underlying Shares Meshblock	-.088 (.005)***		-.048 (.006)***	-.183 (.026)***		-.163 (.026)***
Basic + All Fragment + Underlying Shares Area Unit		-.080 (.004)***	-.079 (.005)***		-.181 (.022)***	.123 (.010)***

Note: ***,**, * represent the levels of statistical significance of 1%, 5%, and 10% respectively. Robust standard errors in parentheses.

[a] Obs.=37942 Meshblocks and 2382 Area Units, as the Crime variable is available only for 2001 and 2006.

Table 4 (Cont'd):
Adding covariates as a robustness check on the effect of heterogeneity on volunteering rates (pooled OLS, meshblock N = 49,600, area unit N = 3504)

Specification	(7)	(8)	(9)	(10)	(11)	(12)
	Birthplace Fragment Meshblock	Birthplace Fragment Area Unit	Birthplace Fragment Hybrid	H Income Fragment Meshblock	H Income Fragment Area Unit	H Income Fragment Hybrid
Basic (Same as Table 3) Meshblock	-.103 (.005)***		-.045 (.006)***	-.019 (.004)***		-.003 (.005)
Basic (Same as Table 3) Area Unit		-.034 (.004)***	-.094 (.005)***		-.108 (.004)***	-.073 (.010)***
Basic + Religious Affiliation (ChrSh, OthSh) Meshblock	-.091 (.006)***		-.048 (.006)***	-.009 (.004)**		.003 (.005)
Basic + Religious Affiliation (ChrSh, OthSh) Area Unit		-.016 (.004)***	-.069 (.005)***		-.060 (.004)***	-.060 (.010)***
Basic + Deprivation (OwnHmSh, MedBedrm, Crime)[a] Meshblock	-.097 (.006)***		-.047 (.007)***	-.042 (.005)***		-.024 (.006)***
Basic + Deprivation (OwnHmSh, MedBedrm, Crime)[a] Area Unit		-.032 (.004)***	-.082 (.006)***		-.138 (.005)***	-.089 (.012)***
Basic + Employment (UnempSh, EmpFTSh, NotLFSh) Meshblock	-.098 (.006)***		-.048 (.006)***	-.011 (.005)**		.002 (.005)
Basic + Employment (UnempSh, EmpFTSh, NotLFSh) Area Unit		-.003 (.004)	-.080 (.005)***		-.068 (.004)***	-.066 (.010)***
Basic + Education Levels (NoQualSh, BHSh) Meshblock	-.166 (.006)***		-.073 (.006)***	-.029 (.004)***		-.009 (.005)*
Basic + Education Levels (NoQualSh, BHSh) Area Unit		-.081 (.004)***	-.167 (.005)***		-.112 (.004)***	-.096 (.010)***
Basic + All Fragment + Underlying Shares Meshblock	.023 (.008)***		.045 (.008)***	.008 (.004)*		.016 (.005)***
Basic + All Fragment + Underlying Shares Area Unit		.089 (.005)***	-.108 (.007)***		-.049 (.004)***	-.015 (.010)

Note: ***,**, * represent the levels of statistical significance of 1%, 5%, and 10% respectively. Robust standard errors in parentheses.

[a] Obs.=37942 Meshblocks and 2382 Area Units, as the Crime variable is available only for 2001 and 2006.

Table 5:
Fixed effects baseline estimation of effects of fragmentation on volunteering rates (meshblock N = 49,600, area unit N = 3504).

Specification (Basic)	(1)	(2)	(3)	(4)	(5)	(6)
	Ethnic Fragment Meshblock	Ethnic Fragment Area Unit	Ethnic Fragment Hybrid	Language Fragment Meshblock	Language Fragment Area Unit	Language Fragment Hybrid
Fragmentation Meshblock	-.046 (.006)***		-.044 (.009)***	.019 (.021)		.025 (.030)
Fragmentation Area Unit		-.112 (.004)***	-.013 (.018)		-.211 (.013)***	-.015 (.025)

Specification (Basic)	(7)	(8)	(9)	(10)	(11)	(12)
	Birthplace Fragment Meshblock	Birthplace Fragment Area Unit	Birthplace Fragment Hybrid	HH Income Fragment Meshblock	HH Income Fragment Area Unit	HH Income Fragment Hybrid
Fragmentation Meshblock	-.033 (.009)***		.021 (.015)	.007 (.005)		.006 (.009)
Fragmentation Area Unit		.001 (.002)	-.081 (.022)***		-.001 (.001)	.001 (.016)

Note: $^{***, **, *}$ represent the levels of statistical significance of 1%, 5%, and 10% respectively.

Table 6:
Summary regarding the (lower bound of) total effects of heterogeneity on volunteering rates

	(1)	(2)	(3)	(4)
	Baseline Cross Section	Extended Cross Section 1 Fragmentation At a Time	Extended Cross Section All Fragmentations Simultaneously	Baseline Fixed Effects
Ethnic/Racial	-	-	-	-
Language	-	-	-	0
Birthplace	-	-	+	-
Nominal HH Income	-	-	+	0[a]

[a] Becomes '-' if 1996 observations dropped. See footnote 9.

Acknowledgments

The authors would like to thank Rachel Croson, John Helliwell, Janet Johnson, Sherry Xin Li, Andrea Menclova, Bob Reed, Eik Sweek, and Samuel Thornton for their helpful comments, as well as the discussant and participants at the 2009 Canadian Economics Association meetings in Toronto and 2010 American Economics Association meetings in Atlanta, and seminar participants at Seoul National University. Funding from the College of Business and Economics at the University of Canterbury, and the Department of Economics at Seoul National University, is gratefully acknowledged.

References

Alesina, A., La Ferrara, E. 2000. Participation in heterogeneous communities. The Quarterly Journal of Economics August, 847-904.

Alesina, A., La Ferrara, E., 2002. Who trusts others? Journal of Public Economics 85, 207-234.

Alesinea, A., La Ferrara, E., 2005. Ethnic diversity and economic performance. Journal of Economic Literature 43, 762-800.

Blakely, T. A., Lochner, K., Kawachi, I., 2002. Metropolitan area income inequality and self-rated health – a multilevel study. Social Science and Medicine 54, 65-77.

Brewster, K. L., 1994. Race difference in sexual activity among adolescent women: the role of neighborhood characteristics. American Sociological Review 59, 408-424.

Costa, D. L., Kahn, M. E., 2003a. Understanding the American decline in social capital, 1952-1998. Kyklos 56, 17-46.

Costa, D. L., Kahn, M. E., 2003b. Civic engagement and community heterogeneity: an economist's perspective. Perspectives on Politics 1, 103-111.

Deaton, A., 2003. Health, inequality, and economic development. Journal of Economic Literature 41, 113-158.

Deaton, A., Lubotsky, D., 2003. Mortality, inequality and race in American cities and states. Social Science and Medicine 56, 1139-1153.

Durlauf, S. N., 2004. Neighborhood effects, in Henderson, J.V., Thisse J.F. (Eds.) Handbook of Regional and Urban Economics, Volume 4. Elsevier. pp. 2175-2242.

Gustavsson, M., Jordahl, H., 2008. Inequality and trust in Sweden: Some inequalities are more harmful than others. Journal of Public Economics 92, 348-365.

Handy, F., Srinivasan, N., 2005. The demand for volunteer labor: A study of hospital volunteers. Nonprofit and Voluntary Sector Quarterly 34, 491- 509.

Hansmann, H., Quigley, J., 1982. Population heterogeneity and the sociogenesis of homicide. Social Forces 61, 206-224.

Habyarimana, J., Humphreys, M., Posner, D., Weinstein, J., 2007. Why does ethnic diversity undermine public goods provision? American Political Science Review 101, 709-725.

Leigh, A., 2006. Trust, inequality and ethnic heterogeneity. The Economic Record 82, 268-280.

Letki, N., 2008. Does diversity erode social cohesion? Social capital and race in British neighborhoods. Political Studies 56, 99-126.

Lochner, K., Pamuk, E. R., Makuc, D., Kennedy, B.P., Kawachi, I., 2001. State level income inequality and individual mortality risk: a prospective multilevel study. Amercian Journal of Public Health 91, 385-91.

Luttmer, E. F., 2001. Group loyalty and the taste for redistribution. Journal of Political Economy 109, 500-528.

Lynch, J. W., Kaplan, G.A., Pamuk, E.R., Cohen, R.D., Heck, K.E., Balfour, J.L., Yen, I.H., 1998. Income inequality and mortality in metropolitan areas of the United States. American Journal of Public Health 88, 1074-80.

Mellor, J. M., Milyo, J., 2001. Re-examining the evidence of an ecological association between income inequality and health. Journal of Health, Politics, Policy and Law 26, 487-522

Mellor, J. M., Milyo, J., 2002. Income inequality and individual health: evidence from the current population survey. Journal of Human Resources 37, 510-39.

Miguel, E., Gugerty, M.K., 2005. Ethnic diversity, social sanctions, and public goods in Kenya. Journal of Public Economics 89, 2325-2368.

Phillips, J., 2008. History of immigration. Te Ara - the Encyclopedia of New Zealand, www.teara.govt.nz/NewZealanders/NewZealandPeoples/HistoryOfImmigration/en , updated November 18th, 2008.

Putnam, R. D., 2007. E pluribus unnum: diversity and community in the Twenty-first Century. The 2006 Johan Skytte prize lecture. Scandinavian Political Studies 30, 137-174.

Rappaport, J., 1999. Local growth empirics. CID Working Paper, Harvard University, No. 23.

Tiebout, C. M., 1956. A pure theory of local expenditures. The Journal of Political Economy 64, 416-424.

Vigdor, J., 2004. Community composition and collective action: Analyzing initial mail responses to the 2000 census. The Review of Economics and Statistics 86, 303-312.

Not for Publication

Appendix 1: Description of Variables

Variable	Description
Volunteering	
Volunteering *VolNarr06* *VolNarr01* *VolNarrAAlt96*	2006, 2001 Proportion of meshblock reporting "Other Helping or Voluntary Work For or Through any Organisation, Group or Marae" in the previous four weeks. Excludes following unpaid activities outside the household: caring for a child or someone who is ill, elderly, or disabled. Construction: "Other Helping…."/(Total – Not Stated) Assumes: "Not Stated" identical in likelihood of volunteering as those who state. 1996 Proportion of meshblock reporting "Attended Committee Meeting etc Unpaid for Group, Church or Marae" in the previous four weeks. Construction: "Attended…"/(Total – Not Specified) Assumes: "Not Specified" identical in likelihood of volunteering as those who state. Excludes the non mutually exclusive categories of "Did Unpaid Training, Coaching, Teaching etc." and "Did Fundraising, Selling etc Unpaid for Group, Church or Marae" and "Did Other Unpaid Work". This was because the latter categories had massive censoring, and had such overlap with the included category that retaining them would have resulted in implausibly high volunteering rates in comparison to 2001 and 2006.
Heterogeneity Measures	
Ethnic/Racial Fragmentation *EthFrag06* *EthFrag01* *EthFrag96*	2006,2001,1996. A fragmentation index for each meshblock, where the five possible ethnic shares s_i are "European", "Maori", "Pacific Peoples", "Asian", and "Middle Eastern/Latin American/African". Individuals could select more than one ethnicity, so the ethnic share is calculated from a baseline of total ethnic affiliations rather than total people. Construction: $1-\sum_{i=1}^{5} s_i^2$ See construction of ethnic shares

Not for Publication

Appendix 1 (Cont'd): Description of Variables

Variable	Description

Heterogeneity Measures (Cont'd)

Language Fragmentation *LanFrag06* *LanFrag01* *LanFrag96*	2006, 2001, 1996. A fragmentation index for each meshblock, where the four possible language shares s_i are "English", "Maori", "Samoan" and "Other". Individuals could select more than one language spoken, so the language share is calculated from a baseline of total language responses rather than total people. Construction: $1-\sum_{i=1}^{4} s_i^2$ See construction of language shares.
Birthplace Fragmentation *BornFrag06* *BornFrag01* *BornFrag96*	2006, 2001, 1996. A fragmentation index for each meshblock's usually resident population, where the two possible shares s_i of each meshblock are "Born in New Zealand" and "Born Overseas". Construction: $1-\sum_{i=1}^{2} s_i^2$ See construction of birthplace shares.
Household Income Fragmentation *HHIncFrag06* *HHIncFrag01* *HHIncFrag96*	2006, 2001, 1996. A fragmentation index constructed over the shares s_i of a neighborhood's households with income from all sources in one of six nominal income bands. The bands are not adjusted for inflation, and are \$0-\$20,000; \$20,001 - \$30,000; \$30,001 - \$50,000; \$50,001 - \$70,000; \$70,000 - \$100,000; \$100,001 and greater. Construction: $1-\sum_{i=1}^{6} s_i^2$ Shares constructed from the frequency of households reporting total income within a band, divided by all households for whom total income was available.

Not for Publication

Appendix 1 (Cont'd): Description of Variables

Variable	Description
Control Variables	
Ethnic Shares *EthEurSh06,01,96* *EthMaoSh06,01,96* *EthPacSh06,01,96* *EthAsnSh06,01,96* *EthMELAA06,01,96*	2006,2001,1996. The proportion of meshblock usual residents reporting one of five ethnic identifications: European, Maori, Pacific, Asian, and Middle Eastern/Latin American/African. Individuals could select more than one ethnicity, so each ethnic share is calculated from a base of the total ethnic affiliations across these five categories rather than total people. Construction: frequencies were summed across the five categories to create a base of total ethnic affiliations from which shares were calculated. For 1996 and 2001, the very small fraction of individuals with "other" ethnicities, such as North American Inuit or Indian, Mauritian, etc. are excluded from the baseline. Statistics NZ assigned the small fraction answering "New Zealander" in 1996 and 2001 as European. For 2006, a much larger proportion of respondents replied "New Zealander", and though 90% of these are thought to be European, they were classified by Statistics NZ under "other." Because "New Zealander" responses made up over 99% of "other" in 2006, we assigned the "other" category as European for that year.
Language Shares *EngLanSh06,01,96* *MaoLanSh06,01,96* *SamLanSh06,01,96* *OthLanSh06,01,96*	2006, 2001, 1996. The proportion of meshblock usual residents indicating they spoke one of four language classifications: English, Maori, Samoan and Other. Individuals could select more than one language (or none), so the language share is calculated from a baseline of total languages spoken rather than total people. Construction: frequencies were summed across the four language categories, omitting "None" or "Not Elsewhere Included", to create a base of total meshblock languages spoken from which shares were calculated.
Birthplace Shares *NZBornSh06* *NZBornSh01* *NZBornSh96*	2006, 2001, 1996. The proportion of meshblock usual residents born in New Zealand vs. born overseas. Construction: frequences were summied across the two birthplace categories, excluding those "Not Elsewhere Specified". Assumed: that those who did not answer this question were as likely to be born overseas as those who did answer.

Not for Publication

Appendix 1 (Cont'd): Description of Variables

Variable	Description
Control Variables (Cont'd)	
Real Median Household Income *RHHIncMed06* *RHHIncMed01* *RHHIncMed96*	2006, 2001, 1996. The median household income from all sources for usual residents of meshblock aged 15 or older. Provided by Statistics New Zealand. Deflated by GDP deflator (1995 = 1000) of 1996 (1016.00), 2001 (1103.50) and 2006 (1224.50).
Female *FemaleSh06* *FemaleSh01* *FemaleSh96*	2006, 2001, 1996. The proportion of a meshblock's usually resident population that is female. Construction: frequencies were summed across the two categories of "Male" and "Female" to create a base from which shares were calculated. Assumes: sex frequencies are more reliable than the "totals" with rounding provided by Stats NZ.
Number of Residents *UsRes06,01,96*	2006, 2001, 1996. Size of meshblock in terms of usually resident population. Only needed if we try weighted least squares to weight meshblock observations by population size.
Population Density *PopDens06*	2006 only. Census meshblock usually resident population divided by meshblock square kilometers.
Median Age *AgeMed06,01,96*	2006, 2001, 1996. Median age of meshblock usually resident population.
Marital Status *MarrSh06* *MarrSh01* *MarrSh96*	2006, 2001, 1996. The share of each meshblock's usually resident population 15 and over who were currently married, as opposed to 1) never married or 2) separated/ divorced/widowed or 3) who did not answer the question. Construction: the four categories were summed to calculate the base. This assumes that all non-responders are not married

Not for Publication
Appendix 1 (Cont'd): Description of Variables

Variable	Description
Control Variables (Cont'd)	
Crime *Crime06* *Crime01*	2006, 2001. The number of recorded offences per capita for each of the 43 Police Areas in New Zealand. Construction: Statistics New Zealand map all meshblocks into one of 43 Police Areas for which per capita offences data is released.
Family Type *CoupNKSh06,01,96* *CoupKSh06,01,96* *SinParSh06,01,96*	2006, 2001, 1996. The share of meshblock families in private dwellings of three possible types: couples without children, couples with children, and single parent families. Construction: frequencies were summed across the three possible categories to provide a baseline.
Religious Affiliation *ChrSh06,01,96* *NoRel06,01,96* *OthRSh06,01,96*	2006, 2001, 1996. The share of each meshblock's usually resident population identifying with one of three categories: Christian, No Religion and Other Religion. For 2001 and 2006 individuals could identify with more than one religion, so that the base is calculated from total religious affiliations, rather than total people. Construction: Other Religion summed frequencies across Buddhist, Hindu, Islam/Muslim, Judaism, Maori Christian, Spiritualist/New Age and Other Religions. "Not Elsewhere Included" are excluded from the base, which assumes that non-responders are similar to responders.
Education High *BHSh06* *BHSh01* *BHSh96*	2006, 2001, 1996. The share of each meshblock's usually resident population 15 or over whose highest degree is a bachelor's or honours degree. Construction: summed frequencies of "Bachelor's Degree or Level 7 Qualification" and "Postgraduate and Honours Degrees" (which excludes masters and PhD degrees), and divided by total people. This assumes that all "Not Elsewhere Included" individuals do not have a bachelor's or honour's degree.

Appendix 1 (Cont'd): Description of Variables

Variable	Description
Control Variables (Cont'd)	
Education Low *NoQualSh06* *NoQualSh01* *NoQualSh96*	2006, 2001, 1996. The share of each meshblock's usually resident population 15 or over who left high school without any (even minimum) qualification. Construction: "No Qualification" divided by total people. This assumes that all "Not Elsewhere Included" individuals had one of the other eight sub-university or four university level degrees.
Mean Household Size *HHSize06* *HHSize01* *HHSize96*	2006, 2001, 1996. The average number of usually resident people per household in the meshblock. Used as a proxy for household crowding and neighborhood deprivation. Construction: provided directly from Statistics New Zealand to zero decimal places.
Labor Force Status *EmpFTSh06,01,96* *EmpPTSh06,01,96* *UnempSh06,01,96* *NotLFSh06,01,96*	2006, 2001, 1996. The share of the usually resident population in each meshblock aged 15 or over in one of four possible categories of labor force status: employed full time, employed part-time, unemployed, or not in labor force. Construction: frequencies for four categories summed to provide a baseline from which shares calculated. "Status Unidentifiable" were excluded, which assumes that those who did not disclose their labor force status were similar to those who did.
Number of Bedrooms *MedBedrms06* *MedBedrms01* *MedBedrms96*	2006, 2001, 1996. Median number of bedrooms in privately occupied dwellings in meshblock. Another proxy for neighborhood deprivation. Construction: provided directly by Statistics New Zealand.

Not for Publication

Appendix A (Cont'd): Description of Variables

Variable	Description
Control Variables (Cont'd)	
Home Ownership Status *OwnHmSh01* *OwnHmSh96*	2001, 1996. The share of dwellings owned or partially owned by their usual residents. Excludes from consideration residents who owned or partially owned their own homes via family trusts. Construction: frequencies for dwellings 1) owned/partially owned by residents, and 2) not owned by residents summed to provide a base from which the share of owner occupied dwellings calculated. Excludes "Dwellings Held in a Family Trust" and "Not Elsewhere Included"
Home Ownership Status *AltOwnHmSh06*	2006 only. The share of dwellings owned or partially owned by their usual residents, or held in a family trust. Dwellings held in a family trust are treated as owned/partially owned. Construction: frequencies for dwellings 1) owned/partially owned by residents, 2) not owned by residents and 3) held in family trusts, summed to provide a base from which the share of owner/trust occupied dwellings calculated. Excludes "Not Elsewhere Included", which assumes non responders are similar in distribution to responders.
Receiving Domestic Purposes Benefit *DomBenSh06* *DomBenSh01* *DomBenSh96*	2006, 2001, 1996. Share of meshblock individuals aged 15 or over receiving the Domestic Purposes Benefit (a welfare programme for single parents). Another proxy for neighborhood deprivation. Construction: frequency of individuals 15 or over receiving income from the domestic purposes benefit in meshblock divided by the total number of people who disclosed their sources of income. This assumes that the distribution of Benefit recipients similar among those who did and did not disclose their sources of personal income.

Not for Publication

Appendix B: Descriptive Statistics (At Meshblock Level, Maximum Sample)

Variable	Obs	Simple Mean	Weighted Mean	Simple Std. Dev.	Min	Max
Volunteering						
VolNarrAAlt96	16712	.1943	.1889	.0702	.0195	.75
VolNarr01	32888	.1682	.1625	.0796	0	1
VolNarr06	34710	.1621	.1540	.0819	0	1
Heterogeneity						
Ethnicity/Race Fragmentation (5 categories)						
EthFrag96	34563	.2955	.3179	.1877	0	.7778
EthFrag01	35150	.3115	.3367	.1932	0	.7951
EthFrag06	34089	.3393	.3629	.1916	0	.7937
Language Fragmentation (4 categories)						
LanFrag96	34791	.2044	.2213	.1414	0	.685
LanFrag01	35323	.2209	.2406	.1449	0	.72
LanFrag06	35975	.2353	.2587	.1467	0	.6837
Birthplace Fragmentation (2 categories)						
BornFrag96	36533	.2389	.2634	.1424	0	.5
BornFrag01	37057	.2509	.2797	.1463	0	.5
BornFrag06	37637	.2754	.3089	.1474	0	.5
Household Income Fragmentation (6 categories)						
HHIncFrag96	25582	.7346	.7421	.0809	0	.8333
HHIncFrag01	26534	.7478	.7556	.0744	0	.8333
HHIncFrag06	27576	.7589	.7652	.0714	0	.8333
Controls for Neighborhood Characteristics						
Ethnic Shares						
EthEurSh96	34563	.7803	.7643	.1976	0	1
EthEurSh01	35150	.7615	.7423	.2108	0	1
EthEurSh06	34089	.7362	.7148	.2180	0	1
EthMaoSh96	34563	.1323	.1333	.1438	0	1
EthMaoSh01	35150	.1327	.1313	.1466	0	1
EthMaoSh06	34089	.1337	.1292	.1420	0	1

Not for Publication

Appendix B (Cont'd): Descriptive Statistics (At Meshblock Level, Maximum Sample)

Variable	Obs	Simple Mean	Weighted Mean	Simple Std. Dev.	Min	Max
Controls for Neighborhood Characteristics (Cont'd)						
Ethnic Shares (Cont'd)						
EthPacSh96	34563	.0438	.0521	.0992	0	1
EthPacSh01	35150	.0489	.0583	.1105	0	1
EthPacSh06	34089	.0533	.0622	.1140	0	1
EthAsnSh96	34563	.0399	.0462	.0695	0	.8857
EthAsnSh01	35150	.0517	.0619	.0865	0	1
EthAsnSh06	34089	.0697	.0855	.1100	0	.925
EthMELAASh96	34563	.0036	.0040	.0143	0	.8
EthMELAASh01	35150	.0052	.0062	.0174	0	.52
EthMELAASh06	34089	.0071	.0083	.0198	0	.52
Language Shares						
EngLanSh96	34791	.8772	.8675	.0976	.2857	1
EngLanSh01	35323	.8658	.8537	.1016	.2222	1
EngLanSh06	35975	.8549	.8393	.1055	.4	1
MaoLanSh96	34791	.0387	.0385	.0595	0	.5714
MaoLanSh01	35323	.0391	.0382	.0579	0	.5172
MaoLanSh06	35975	.0369	.0347	.0571	0	.5
SamLanSh96	34791	.0138	.0168	.0401	0	.5
SamLanSh01	35323	.0147	.0179	.0406	0	.4444
SamLanSh06	35975	.0143	.0174	.0395	0	.4444
OthLanSh96	34791	.0703	.0772	.0690	0	.5714
OthLanSh01	35323	.0803	.0903	.0769	0	.7778
OthLanSh06	35975	.0939	.1086	.0866	0	.6
Born in New Zealand Shares						
NZBornSh96	36533	.8402	.8243	.1218	0	1
NZBornSh01	37057	.8260	.8050	.1353	0	1
NZBornSh06	37637	.7998	.7703	.1498	0	1

Not for Publication

Appendix B (Cont'd): Descriptive Statistics (At Meshblock Level, Maximum Sample)

Variable	Obs	Simple Mean	Weighted Mean	Simple Std. Dev.	Min	Max
Controls for Neighborhood Characteristics (Cont'd)						
Real Median Household Meshblock Income (1995 New Zealand Dollars)						
RHHIncMed96	25587	37307	37768	14760.5	0	98425
RHHIncMed01	26534	38767	39328	15965.8	7250	90621
RHHIncMed06	27576	44333	45276	17038.2	2858	81666
Sex						
FemaleSh96	37103	.5015	.5088	.0781	0	1
FemaleSh01	37544	.5051	.5123	.0770	0	1
FemaleSh06	38090	.5052	.5121	.0741	0	1
Population Density of Meshblock 2006						
PopDens06	41362	1783.5	2503.3	2595.6	0	143775
Median Age						
AgeMed96	31903	33.65	33.44	8.0842	10	86
AgeMed01	32471	35.42	35.16	8.3549	11	88
AgeMed06	33210	36.60	36.22	8.7959	13	88
Share Married (Of Age 15 or Older)						
MarrSh96	35542	.4936	.4858	.1649	0	1
MarrSh01	36068	.4707	.4654	.1633	0	1
MarrSh06	36812	.4513	.4879	.1627	0	1
Per Capital Recorded Offences						
Crime01	41376	.1115	.1099	.0699	.0689	.7073
Crime06	41376	.1009	.1012	.0436	.0641	.4178
Family Type Shares						
CoupleNKSh96	32340	.3757	.3680	.1631	0	1
CoupleNKSh01	32875	.3924	.3833	.1663	0	1
CoupleNKSh06	33660	.4034	.3931	.1698	0	1

Not for Publication

Appendix B (Cont'd): Descriptive Statistics (At Meshblock Level, Maximum Sample)

Variable	Obs	Simple Mean	Weighted Mean	Simple Std. Dev.	Min	Max
Controls for Neighborhood Characteristics (Cont'd)						
Family Type Shares (Cont'd)						
CoupleKSh96	32340	.4507	.4505	.1586	0	1
CoupleKSh01	32875	.4202	.4225	.1552	0	1
CoupleKSh06	33660	.4157	.4204	.1542	0	1
SingleParSh96	32340	.1736	.1815	.1433	0	1
SingleParSh01	32875	.1874	.1942	.1440	0	1
SingleParSh06	33660	.1808	.1865	.1423	0	1
Religious Affiliation Shares						
ChrisSh96	33546	.6826	.6794	.1243	0	1
ChrisSh01	34123	.6251	.6215	.1278	0	1
ChrisSh06	34763	.5664	.5634	.1304	0	1
NoRelSh96	33546	.2760	.2761	.1139	0	1
NoRelSh01	34123	.3150	.3140	.1177	0	1
NoRelSh06	34763	.3651	.3613	.1262	0	1
OthRSh96	33546	.0415	.0445	.0630	0	.9565
OthRSh01	34123	.0599	.0644	.0752	0	1
OthRSh06	34763	.0685	.0753	.0842	0	.9583
Education Shares High or Low						
Bach/HonsSh96	31128	.0777	.0784	.0861	0	.7778
Bach/HonsSh01	31824	.0974	.0990	.0967	0	.75
Bach/HonsSh06	31646	.1134	.1162	.0913	0	.7
NoQualSh96	31128	.3301	.3265	.1384	0	.9512
NoQualSh01	31824	.2455	.2405	.1158	0	.96
NoQualSh06	31646	.2335	.2264	.1157	0	.8313
Average Household Size						
HHSize96	27936	3.0056	2.8102	2.4534	1	35
HHSize01	28824	2.9295	2.7418	2.4105	1	37
HHSize06	29807	2.9289	2.7613	2.3576	1	39

Not for Publication

Appendix B (Cont'd): Descriptive Statistics (At Meshblock Level, Maximum Sample)

Variable	Obs	Simple Mean	Weighted Mean	Simple Std. Dev.	Min	Max
Controls for Neighborhood Characteristics (Cont'd)						
Labor Market Shares						
EmplFTSh96	34989	.4741	.4640	.14007	0	1
EmplFTSh01	35559	.4847	.4747	.13909	0	1
EmplFTSh06	36263	.5113	.5023	.13149	0	1
EmplPTSh96	34989	.1419	.1398	.0671	0	1
EmplPTSh01	35559	.1444	.1424	.0652	0	.6667
EmplPTSh06	36263	.1504	.1486	.0652	0	1
UnempSh96	34989	.0498	.0521	.0520	0	.5
UnempSh01	35559	.0492	.0513	.0504	0	.5
UnempSh06	36263	.0339	.0357	.0403	0	1
NotLFSh96	34989	.3342	.3441	.1414	0	1
NotLFSh01	35559	.3218	.3315	.1370	0	1
NotLFSh06	36263	.3044	.3134	.1304	0	1
Median Number of Bedrooms						
MedBedrms96	26867	2.9106	2.913	.4489	0	6
MedBedrms01	27868	2.9693	2.974	.4742	1	6
MedBedrms06	28885	3.0010	3.010	.5031	1	6
Share Owning or Partially Owning Own Home						
OwnHmSh96	33809	.7002	.7049	.2100	0	1
OwnHmSh01	34511	.6796	.6754	.2025	0	1
AltOwnHmSh06	34106	.6619	.6624	.2046	0	1
Share Receiving Domestic Purposes Benefit						
DomBenSh96	29504	.0416	.0427	.0472	0	.4545
DomBenSh01	30153	.0414	.0423	.0453	0	.4286
DomBenSh06	31110	.0339	.0342	.0407	0	.4444

Not for Publication

Appendix C: Population Weighted Means and Standard Deviations of Key Variables Over Time (NZ Census, 1996, 2001, 2006), at the Meshblock and Area Unit Levels, using Maximum Sample.

Variable	*Neighbor-hood Size*	*Census Year*					
		1996		2001		2006	
		Mean (St. Dev)	N	Mean (St. Dev)	N	Mean (St. Dev)	N
Volunteering Rate	Meshblock	.189 (.066)	16712	.163 (.069)	32888	.154 (.068)	34710
	Area Unit	.193 (.044)	1757	.163 (.042)	1779	.154 (.039)	1780
Ethnic/Racial Fragmentation	Meshblock	.318 (.184)	34563	.337 (.190)	35150	.363 (.190)	34089
	Area Unit	.342 (.161)	1786	.362 (.170)	1781	.387 (.173)	1772
Language Fragmentation	Meshblock	.221 (.137)	34791	.241 (.140)	35323	.259 (.143)	35975
	Area Unit	.231 (.114)	1791	.250 (.120)	1784	.268 (.124)	1787
Birthplace Fragmentation	Meshblock	.263 (.130)	36533	.280 (.134)	37057	.309 (.132)	37637
	Area Unit	.271 (.112)	1806	.289 (.118)	1798	.319 (.117)	1803
Household Income Fragmentation	Meshblock	.742 (.075)	25582	.756 (.069)	26534	.765 (.067)	27576
	Area Unit	.782 (.039)	1748	.796 (.033)	1755	.805 (.033)	1753