Investigating a wage curve for New Zealand

Kerry L. Papps*+

This paper examines evidence for a stable inverse relationship between the wages paid to workers and the unemployment rate across local labour markets in New Zealand, a phenomenon known as the wage curve. A variety of specifications of the wage curve are examined. Overall, weighted least squares estimates reveal a value of the unemployment elasticity of pay that is close to the international consensus estimate of –0.1. Some support is also found for the concept of a positive long-run relationship between wages and unemployment existing alongside the wage curve. However, there is evidence of potential endogeneity of the unemployment rate, although data limitations severely restrict the availability of suitable instruments.

“Science is organised common sense where many a beautiful theory was killed by an ugly fact.”

THOMAS HUXLEY

1. Introduction

It is becoming increasingly obvious that the solutions to a number of important social and economic issues lie in a thorough understanding of the nature and causes of inter-regional dispersion in labour market outcomes. These include the most appropriate policy for improving the labour market outcomes of depressed areas and whether such a policy should be implemented at a national level or targeted at specific regions or groups of workers. However, macroeconomic models have traditionally neglected the effects of spatial variation in labour market attributes among individuals and jobs, instead assuming the existence of a single representative labour market for all workers of a given description. Such a simplification is unable to provide any explanation for the existence of geographic differentials in employment conditions.

In their 1994 book, Blanchflower and Oswald (1994) provided an important contribution to the literature on local labour markets. Drawing on individual level data

* Ph.D. student, Department of Economics, Cornell University.

+ This paper is based on research that was undertaken for a Master of Commerce and Administration thesis at Victoria University of Wellington. In this respect, the author would like to acknowledge the invaluable contributions that were made by Associate Professors Jacques Poot and Philip Morrison. Useful comments on the current paper were also made by Sylvia Dixon and by an anonymous referee. The author was awarded the 2001 A.R. Bergstrom Prize in Econometrics for this paper.
from twelve countries and the findings of other studies, the authors documented evidence of a stable downward-sloping convex curve linking local unemployment and the level of pay. This relationship, which Blanchflower and Oswald christened the “wage curve”, was found to roughly take the form: $\ln w = -0.1 \ln u + \text{other terms}$, where $w$ denotes the real wage of an individual or group and $u$ the prevailing unemployment rate in the local labour market.

Blanchflower and Oswald noted the “remarkable” uniformity of the results of their and others’ wage curve studies across different nations and concluded that “every country seems to have a ‘wage curve’” (p. 12). They believed that the wage curve is not a mislabelled labour supply curve and is incompatible with the competitive model of the labour market. Instead, they offered three non-competitive explanations: a labour contract model, an efficiency wage framework and a union bargaining model.\footnote{The authors favoured the latter two models, as they are consistent with other patterns that were discovered in the data. Papps (2000) and Morrison et al. (2000) preferred an efficiency wage explanation for the wage curve that differs from that of Blanchflower and Oswald in that it involves firms minimising turnover costs.} Subsequent investigations have largely supported this finding, such as Janssens and Konings (1998) and Baltagi et al. (2000).\footnote{In a recent meta-analysis of studies of the relationship between wages and unemployment, Nijkamp and Poot (2000) concluded that the wage curve is a robust empirical phenomenon, despite evidence of publication bias.}

The existence of a wage curve appears to contradict an earlier view advanced by Harris and Todaro (1970) and supported by the work of Hall (1970; 1972) and Reza (1978). Invoking Adam Smith’s (1776) notion of compensating wage differentials, Harris and Todaro argued that migration of labour force participants takes place until the expected wage is the same in each region.\footnote{In reference to an earlier version of the Harris-Todaro model, Greenwood (1975) noted that while it “is applied specifically to rural-urban migration in less-developed countries, the model is general enough to apply to interregional migration in any country” (p. 403).} The expected wage is the wage multiplied by the probability of getting a job, hence workers in those regions with high risks of being unemployed must be compensated by receiving higher levels of pay and vice versa. A positive wage-unemployment locus across regions is generated as a result.

The essential difference between the frameworks of Harris and Todaro on the one hand and Blanchflower and Oswald on the other is that the former relies on the existence of a competitive labour market with relatively costless mobility. In contrast,
all explanations for the wage curve assume some form of non-competitive behaviour in the labour market. Nevertheless, both theories may still be relevant, Blanchflower and Oswald argued, if “permanent” wages are positively related to “permanent” unemployment, while movements in actual wages and unemployment exhibit a negative correlation.

To date, only one paper has attempted to determine the effect of unemployment on wages in local labour markets in New Zealand. Morrison and Poot (1999) found evidence of a cross-sectional wage curve based on regional data from the 1996 Census of Population and Dwellings. Their estimates of the unemployment elasticity of pay, $\beta$, were in the range of $-0.07$ to $-0.12$. The authors argued that “a better understanding of the dynamics of local labour markets is an essential requirement for further study of the wage curve” (p. 96).

The aim of this study is to examine whether the results of Morrison and Poot hold when using disaggregated census data. Although these do not provide the individual level information that would ideally be used in wage curve regressions, these “synthetic” microeconomic data are much more representative of the labour market than aggregate regional data. An investigation of the robustness of estimates of the wage curve to controls for simultaneity bias is given, as well as an indication of the degree to which the data used in this study lead to a gain in efficiency over those used by Morrison and Poot.

2. Data

Information on the labour market behaviour of workers aged between 15 and 59 was drawn from the 1986, 1991 and 1996 New Zealand Censi of Population and Dwellings. Observations on median annual income, $Y$, and average weekly hours normally worked, $H$, were obtained for 34,560 population subgroups, $i$, each containing $n_i$ wage and salary earners. Since all theories explain the wage curve as the product of the wage setting process, non-wage and salary earners (for example, self-employed individuals) were excluded. The subgroups form a mutually exclusive

---

4 occupational groups and 39 urban areas were considered, giving a total of 156 observations.

5 For confidentiality reasons, Statistics New Zealand randomly rounds all values of $n_i$ to base 3. A reported count of 6 could, therefore, imply an actual count of anywhere between 4 and 8. Since the long-run expected value is designed to equal the original count, the impact of this on the results of this study should be insignificant.
and exhaustive set of permutations of the following qualitative characteristic categories: year (3 values); urban area (30 values); age (3 values); educational attainment (2 values); ethnicity (2 values); gender (2 values); full-time/part-time status (2 values); occupation (8 values).6

Attention was restricted to workers in 30 urban areas and 8 occupations in order to control for as many unobservable characteristics as possible, in a sense allowing each group to closely “resemble” an individual worker. A drawback of this approach is that it will lead to selectivity bias in the regressions reported later if workers in either omitted locations or omitted occupations have different unobservable characteristics to included workers. This is most likely to occur in the latter case, however the occupation groups were chosen as they are common and, collectively, are broadly representative of the labour market. Nonetheless, any conclusions drawn should only be interpreted as applying to workers in the specific occupations and urban areas chosen.

A large number of subgroups reported a value of zero for \( n \) and, hence, \( Y \) and \( H \), due to the fact that no workers matched the particular description. These observations were deleted, leaving 20,302 observations in total. Collectively, around half a million workers are covered in each year, or approximately 56% of all wage and salary earners in New Zealand.7

The appropriate measure of the wage, \( w \), to include in the wage curve is the real hourly earnings of workers. Without knowing either the annual hours or weeks of work for a subgroup, no value can be derived from the census that corresponds to this measure exactly. One alternative is to use annual earnings as a proxy for \( w \), as Blanchflower and Oswald (1994) did for the United States. However, such an approach neglects the fact that knowledge of \( H \) allows variation in weekly hours worked to be controlled for. Accordingly, the wage level in 1986 dollars of a group \( i \) of gender \( g \) in urban area \( j \) and year \( t \) is calculated as follows, using national Consumers’ Price Index values for each year, \( P \):8

---

6 The education categories identify whether the members of a subgroup have attained a post-secondary school qualification or not, ethnicity refers to Europeans and non-Europeans and the following age groups are used: 15-25, 26-40 and 41-59. Each occupation group comprises a number of related four-digit occupations. These are listed in the appendix, along with the urban areas selected.

7 532,755 wage and salary earners are included in the 1986 sample, 497,136 in 1991 and 565,788 in 1996.

8 Regressions that include time-specific effects automatically control for the impact of \( P \). Nonetheless, it is included in equation (1) for expositional convenience. Ideally, an urban area-specific price level
\[ w_{gij} = \frac{Y_{gij}}{52H_{gij}p_r}. \] (1)

The wage computed by equation (1) is the best available proxy for hourly earnings, using census data. However, it must be noted that this measure assumes that \( H \) is the average hours worked per week over the previous 52 weeks, which differs from the census definition and means that the wages of those who have not worked the entire year will be under-estimated.\(^9\) In addition, since \( Y \) includes non-wage income, the wages of those who receive a large amount of interest, dividends, etc., will be over-estimated.

A variety of aggregate statistics for each urban area were collected from the 1981, 1986, 1991 and 1996 Censuses. These include the number of unemployed males and females and the number of each gender in the labour force. The variables \( u_T, u_M \) and \( u_F \) refer to the total, male and female unemployment rates, respectively, and are constructed by dividing the relevant number of unemployed people by the relevant number in the labour force. The principal unemployment rate used in this paper, \( u \), was constructed so as to be specific to each gender, in addition to each urban area and year, as follows.\(^10\)

\[ u = \begin{cases} 
  u_M & \text{if male} \\
  u_F & \text{if female} 
\end{cases} \] (2)

3. Analysis

Central to Blanchflower and Oswald’s argument for the wage curve is the assumption that the unemployment rate is the appropriate measure of local labour market tightness to be added to an earnings equation. However, a largely agnostic

---

\(^9\) This issue is explored further in Chapter 6 of Papps (2000).

\(^10\) This is consistent with the approach of Kennedy and Borland (2000). It is possible to disaggregate the unemployment rate further, for example allowing it to vary by age group and ethnicity. However, differences in employment opportunities between men and women are arguably the most significant and they represent a clear dichotomy.
view is taken with respect to the actual functional form in which the unemployment rate should be included and a variety of specifications are tested.

Before any conclusion about the role of the local unemployment rate can be reached, it is essential to control for the effect of other relevant determinants of workers’ earnings. Along the lines of Blanchflower and Oswald (1994), the following general form of a function, φ, is posited, determining the natural logarithm of real wages of workers of gender g in subgroup i:

\[
\ln w_{gti} = \phi(x_{gti}, u_{gti}, p_{gti}, \omega_{gti}, \lambda_{gti}).
\]  

(3)

Here, \(u\) denotes the unemployment rate, as defined earlier, \(\omega\) an urban area dummy variable, \(\lambda\) a year dummy variable and \(p\) a relative measure of cross-sectional price dispersion, specifically the average house sale price expressed as a proportion of the corresponding national value.\(^{11}\) \(x\) represents a vector of the relevant characteristics of the particular subgroup of workers, including dummies for age group, post-secondary school qualifications, ethnicity, gender, full-time/part-time status and occupation.

Omitting \(u\) and \(\omega\), equation (3) appears to be a reasonable representation of the earnings function, explaining over 70% of the variation in wages among groups of workers in the sample. Estimates of the wage curve are based on this general specification, which is known as a Mincerian earnings function, after Mincer (1974), and is conventional in the literature on wage determination. As Blanchflower and Oswald (1994) noted, “their novelty is in the inclusion of local unemployment as a regressor” (p. 101). The authors propose a number of specifications of the function \(\phi\) in equation (3), involving the unemployment rate entering in different forms. Their approach was to experiment by estimating each with the data. Table 1 presents the results of adding the unemployment rate, \(u\), to the list of regressors in the earnings equation in some of these forms.\(^{12}\) Each specification generates strong support for a

\[11\] Although \(w\) is constructed by taking account of nationwide changes in consumer prices over time, it does not incorporate cross-sectional variation. As observed by Kennedy and Borland (2000), housing costs are likely to form a significant component of inter-regional differences in the cost of living and are therefore likely to be a good proxy for the latter variable.

\[12\] As with all correlations and regressions referred to later, observations are weighted by the size of the subgroup, \(n_i\), so that values of \(w\) and \(u\) representing the characteristics of a small number of workers are given proportionally less weight than those representing the characteristics of a large group.
downward-sloping wage curve. Following Blanchflower and Oswald’s approach, the five wage curves are graphed in Figure 1 by setting $w$ and $u$ equal to their mean values, which enables the constant to be solved for in each equation. The specifications involve the following regressors:

(i) The natural logarithm of the unemployment rate, $\ln u$.

(ii) The natural logarithm of the unemployment rate, $\ln u$, and its cube, $(\ln u)^3$.

(iii) The unemployment rate, $u$.

(iv) The unemployment rate, $u$, and its square, $u^2$.

(v) The reciprocal of the unemployment rate, $u^{-1}$.

All five specifications result in wage curves that are similar at all but the highest observed unemployment rates. In contrast to Blanchflower and Oswald, the non-linear terms are found to be unimportant, with no evidence of an upward-sloping portion of the wage curve in practice. The estimated coefficient on the cubic term is insignificant in specification (ii) and although the quadratic term in specification (iv) is found to be

$$
\begin{array}{cccccc}
\text{Dependent variable} & \text{Specification} \\
\text{variable} & (i) & (ii) & (iii) & (iv) & (v) \\
\ln u & -0.098^{c} & -0.114^{c} & & & \\
(\ln u)^3 & \ (0.008) & \ (0.030) & & & \\
u & -1.013^{c} & -1.974^{c} & & & \\
(0.086) & (0.227) & & & \\
u^2 & 4.267^{c} & & & & \\
(0.934) & & & & & \\
u^{-1} & \ 0.006^{c} & \ & & & \\
(0.001) & & & & & \\
R^2 & 0.710 & 0.710 & 0.709 & 0.710 & 0.710 \\
\end{array}
$$

Note: Estimated coefficients for regressors other than the unemployment rate have been omitted. Significance at the 1%, 5% and 10% level is denoted by $^c$, $^b$ and $^a$, respectively. Standard errors are in parentheses.

13 Moulton (1986) noted that the inclusion of an explanatory variable in a regression which is defined at a more aggregate level than the dependent variable may lead to $t$-statistics on the aggregate variable that are biased upwards. Specifically, a difficulty will arise here if the earnings of workers of the same gender and in the same urban area share some common component of variation that is not attributable either to measured characteristics or to the unemployment rate, meaning that the error term will be positively correlated between workers in the same local labour market.

14 Each wage curve passes through the intersection of the average values, $\overline{w}$ and $\overline{u}$, by construction. They are plotted only over the range of unemployment rates that are present in the data, which is 0.03 to 0.25.
significant, it is such that the wage curve only has a positive slope above $u = 0.23$.\textsuperscript{15}

“Principally for ease of interpretation and computation” (p. 102), Blanchflower and Oswald eventually settled on specification (i) as their standard specification of the wage curve, so that $\ln u$ is the only variable added to the earnings equation. Since the dependent variable is also in logarithmic form, this implies a constant unemployment elasticity of pay. According to Table 1, a Mincerian earnings equation with $\ln u$ only added is found to minimise both the Akaike (1973) information criterion ($AIC$) and the Schwarz (1978) criterion ($SC$) for New Zealand, therefore the double logarithmic specification was also chosen here.

\textsuperscript{15} Winter-Ebmer (1996) reached a similar conclusion and rejected the notion of a U-shaped wage curve.
The first column of Table 2 reveals the results of estimating the earnings function when ln $u$ is added to the explanatory variables in $x$, ln $p$ and $t$ in equation (3). The estimate of the unemployment elasticity of pay obtained is equal to $-0.098$, remarkably close to the “magic” value of $-0.1$ propounded by Blanchflower and Oswald. This result implies that a 10% increase in an urban area’s gender-specific unemployment rate will induce a 1.0% fall in the real wages of workers in the urban area. Relative to the variables contained in $x$, ln $u$ is not a statistically important determinant of wages and increases the equation’s value of $R^2$ only slightly. All other independent variables enter the earnings equation with coefficients of the expected sign.

The level of pay received by workers in a particular locality may be influenced by other relevant variables that are specific to the local labour market but are relatively

<table>
<thead>
<tr>
<th>Regressor</th>
<th>Specification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>(i) WLS</td>
</tr>
<tr>
<td></td>
<td>1.551 $^c$</td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
</tr>
<tr>
<td>ln $u$</td>
<td>(ii) WLS</td>
</tr>
<tr>
<td></td>
<td>-0.098 $^c$</td>
</tr>
<tr>
<td></td>
<td>(0.008)</td>
</tr>
<tr>
<td>Aged 26-40</td>
<td>(iii) WLS</td>
</tr>
<tr>
<td></td>
<td>0.353 $^c$</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td>Aged 41-60</td>
<td>(iv) IV</td>
</tr>
<tr>
<td></td>
<td>0.392 $^c$</td>
</tr>
<tr>
<td>Qualified</td>
<td>(i) WLS</td>
</tr>
<tr>
<td></td>
<td>0.129 $^c$</td>
</tr>
<tr>
<td>European</td>
<td>(ii) WLS</td>
</tr>
<tr>
<td></td>
<td>0.052 $^c$</td>
</tr>
<tr>
<td>Male</td>
<td>(iii) WLS</td>
</tr>
<tr>
<td></td>
<td>0.174 $^c$</td>
</tr>
<tr>
<td>Full-time</td>
<td>(iv) IV</td>
</tr>
<tr>
<td></td>
<td>0.040 $^c$</td>
</tr>
<tr>
<td>ln $p$</td>
<td>(i) WLS</td>
</tr>
<tr>
<td></td>
<td>-0.044 $^c$</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td>ln $\bar{u}$</td>
<td>(ii) WLS</td>
</tr>
<tr>
<td></td>
<td>0.047 $^c$</td>
</tr>
<tr>
<td></td>
<td>(0.003)</td>
</tr>
<tr>
<td>Occupation dummies</td>
<td>Yes</td>
</tr>
<tr>
<td>Time dummies</td>
<td>Yes</td>
</tr>
<tr>
<td>Urban area dummies</td>
<td>No</td>
</tr>
<tr>
<td>Number of observations</td>
<td>20,302</td>
</tr>
</tbody>
</table>

Note: Significance at the 1%, 5% and 10% level is denoted by $^c$, $^b$ and $^a$, respectively. Standard errors are in parentheses.
constant over time, such as amenities and spatial location. To capture the effect of these “permanent” urban area attributes, a full set of 29 urban area dummies, \( w \), is included in the second column of Table 2. The specification of the earnings equation, \( f \), now estimated is:

\[
\ln w_{ijt} = \gamma_0 + \mathbf{x}_t' \gamma_1 + \gamma_2 \ln p_{jt} + \beta \ln u_{ijt} + \omega_i + \tau_t + \epsilon_{ijt}.
\]

(4)

This is a two-way fixed effects model, with \( \omega \) and \( \tau \) capturing unobservable factors that determine \( \ln w \). Table 3 presents the results of \( F \)-tests for the significance of the urban area and time effects.\(^{16}\) In the first column the two-way fixed effects model is compared with a model with urban area effects only, while in the second it is compared with a model with time effects only. Strong evidence that time and urban area effects are each jointly significant is found, justifying use of a two-way fixed effects model over one which either pools the data or includes only one type of effect.

As the second column of Table 2 indicates, the estimated coefficient on unemployment, \( \hat{\beta} \), now increases in magnitude to \(-0.133\). The exclusion of urban area effects appears to bias upwards the value of \( \hat{\beta} \). This is consistent with

\[ F = \frac{(RSS_R - RSS_U) / s}{RSS_U / (RT - (R - 1) - (T - 1) - k)}, \]

where \( k \) is the number of regressors in the unrestricted model other than dummy variables, \( s \) is the number of degrees of freedom gained from moving from the unrestricted model to the restricted model and \( RSS_U \) and \( RSS_R \) denote the residual sum of squares from the fixed effects regression and the appropriate restricted model, respectively.

\(^{16}\) The general test statistic is given by: \( F = \frac{(RSS_R - RSS_U) / s}{RSS_U / (RT - (R - 1) - (T - 1) - k)} \), where \( k \) is the number of regressors in the unrestricted model other than dummy variables, \( s \) is the number of degrees of freedom gained from moving from the unrestricted model to the restricted model and \( RSS_U \) and \( RSS_R \) denote the residual sum of squares from the fixed effects regression and the appropriate restricted model, respectively.
Blanchflower and Oswald’s results using Current Population Survey data from the United States. They surmised that the mistake made by previous empirical studies that reported a positive wage-unemployment relationship was to omit controls for regional fixed effects. The above results suggest that a wage curve exists regardless of whether \( \omega \) is included or not, although a stronger relationship is found in the former case.\(^{17}\)

If it is possible to model either type of effect as a random effect, the fixed effects estimator, although consistent, is inefficient. Use of the random effects estimator requires the assumption that \( \omega \) and/or \( t \) is uncorrelated with the regressors in the model. The validity of this assumption can be examined using Hausman’s (1978) test. This test involves comparing the estimated parameter values for the explanatory variables under both random and fixed effects specifications. Under the null hypothesis that unemployment is uncorrelated with the random effects, the random effects estimator is efficient. However, if the null hypothesis is false, the use of random effects produces an inconsistent estimator, whereas the fixed effects estimator remains consistent.

The third and fourth columns of Table 3 report the results of tests that the urban area effects are uncorrelated with the independent variables, firstly controlling for time effects under both the null hypothesis and its alternative and then omitting them. Both methods clearly reject the hypothesis of no correlation. Similar results are found when the time effects are tested in the fifth and sixth columns. Hence, the evidence indicates that both types of effects should be captured as dummy variables and that a switch to random effects estimation would introduce bias.

As mentioned in the introduction, Harris and Todaro’s (1970) concept of compensating differentials across regions may coexist with a negatively sloped wage curve if it is taken to refer to the permanent, rather than contemporaneous, values of wages and unemployment. In the regression reported in the second column of Table 2 the estimated values of the urban area effects \( \hat{\omega} \) capture the permanent component of wages in each local labour market in New Zealand over the sample period. Figure 2 plots \( \hat{\omega} \) against the natural logarithm of the time-averaged total unemployment rate for each urban area, \( \ln \bar{\pi}_t \), an estimate of the permanent value of unemployment. The two exhibit a positive relationship, with a significant correlation coefficient of

\(^{17}\) Separate regressions were also estimated for each year and reported evidence of a wage curve in every case, although the unemployment elasticity of pay was found to have decreased in magnitude
Part of this correlation is due to the influence of the outlying urban area, Tokoroa (urban area 17). However, when this observation is omitted, $\omega$ and $\ln T_u$ are still found to have a significant correlation, with a coefficient of 0.698.\[19\]

Figure 2 lends tentative support to Blanchflower and Oswald’s (1995) claim that “movements in actual wages can be negatively correlated with movements in actual unemployment, while at the same time “permanent” unemployment... is positively related to “permanent” wages” (p. 160). Further evidence is found in the third column of Table 2. Here the urban area dummies are replaced by the natural logarithm of the time-averaged unemployment rate, $\ln T_u$. This variable approximates the component of unemployment that is due to the inherent characteristics of different locations, for over the decade.

---

18 This provides an explanation for why the Hausman test rejects the null hypothesis that the urban area effects are uncorrelated with the regressors.

19 More so than any other urban area, Tokoroa depends on pulp and paper processing and is a “one company town”. This monopsony situation means that the urban area dummy for Tokoroa will largely measure the effect of a single industry. If, for example, employees are compensated for unpleasant working conditions, this may explain why the wage level in the community is consistently above the average.
example climate and commuting times. It is found to enter the earnings equation with a positive and significant coefficient. A 10% increase in the current unemployment rate in an urban area leads to a 1.4% fall in wages, while a 10% increase in the long-run expected unemployment rate begets a 0.9% increase in the earnings of workers.\(^20\) Thus, the wage curve is able to be reconciled with the Harris-Todaro notion of compensating differentials across urban areas.

Since it is jointly determined with wages in the labour market, unemployment not only affects pay but is itself a function of the wage rate. Underlying the estimates of the earnings equation obtained thus far is the implicit assumption that, since each observation represents a relatively small group of workers, the (gender-specific) unemployment rate for the urban area as a whole is unaffected by the wage of a particular group and may therefore be considered an exogenous variable. However, if this is not the case, weighted least squares estimates of $\beta$ will be affected by simultaneity bias. As the unemployment rate, $u$, is the average value for an urban area, while the wage, $w$, represents a narrowly defined group of workers, this is not necessarily a problem here, provided that no group dominates their local labour market. Nevertheless, the exogeneity of the unemployment rate was examined by conducting Hausman’s specification error test. This entailed regressing $\ln u$ on all available exogenous variables, extracting the residuals and testing the effect of their inclusion in the earnings equation alongside the other regressors. The residuals from the auxiliary regression were found to have a significant effect on $\ln w$ in the two-way fixed effects model, indicating that the unemployment rate should be considered endogenous.

To overcome simultaneity bias, instrumental variables estimation must be employed. This requires one or more valid instruments, that is, a set of variables that affect unemployment in a region but may be omitted from the earnings equation.\(^21\) Blanchflower and Oswald (1994) noted that in selecting such variables for the earnings equation “there is no well-established literature upon which to draw” (p. 227). Following their suggestions, the following instruments were chosen: the lagged value of $\ln u$, the annual level of rainfall as a deviation from the normal

---

20 Since it is generated from only three observations, $\ln \bar{u}_t$ is a less than ideal measure of permanent unemployment in any region, although the observations are derived from different points on the business cycle.

21 In other words, variables that are correlated with $\ln u$ but not with the error term $\varepsilon$ in equation (4).
amount, $W_R$, the annual sunshine deviation from normal, $W_S$, and an industry mix variable, $I$.\textsuperscript{22} In addition, Oswald (1997) argued that a high rate of house ownership among workers contributes to joblessness by limiting their mobility. Since there is no obvious direct relationship between the proportion of houses in a local labour market that are rented, $r$, and workers’ earnings, $r$ would appear to be another plausible instrument for local unemployment in the earnings equation.

Some evidence of the appropriateness of the various instrumental variables is provided in Table 4. All five are found to have a correlation with $\ln u$ that is significant at the 1% level. High rates of unemployment are associated with low rainfall and sunshine levels, low incidences of renting and low values of the industry mix variable.\textsuperscript{23} In addition, high past values of unemployment tend to coincide with high current values. None of the four contemporaneous variables is correlated with the residuals from weighted least squares estimation of the earnings equation at the 5% level of significance, indicating that they are all valid instruments. The lagged unemployment rate does have a significant correlation coefficient, however its status as a predetermined variable dictates that it too is a valid instrument.\textsuperscript{24}

The final column of Table 2 presents the results of instrumental variables estimation of equation (4), using all five exogenous variables as instruments for $\ln u$. In contrast to Blanchflower and Oswald’s results, once endogeneity of the unemployment rate is controlled for, evidence of a wage curve vanishes. Regardless

\begin{table}[h]
\centering
\caption{Correlations of instrumental variables}
\begin{tabular}{|c|c|c|}
\hline
Instrument & $\ln u$ Correlation & $\varepsilon$ Correlation \\
\hline
$I$ & -0.607 \textsuperscript{c} & -0.004 \\
$W_R$ & -0.022 \textsuperscript{c} & 0.011 \\
$W_S$ & -0.422 \textsuperscript{c} & -0.012 \textsuperscript{a} \\
r & -0.120 \textsuperscript{c} & 0.004 \\
$\ln u_{-1}$ & 0.206 \textsuperscript{c} & -0.019 \textsuperscript{c} \\
\hline
\end{tabular}
\end{table}

Note: Significance at the 1%, 5% and 10% level is denoted by \textsuperscript{c}, \textsuperscript{b} and \textsuperscript{a}, respectively.

\textsuperscript{22} The lagged value of $\ln u$, denoted $\ln u_{-1}$, refers to the corresponding value from the previous census, i.e. a five-year lag. Along the lines of Bartik (1991), $I$ is a weighted average of the growth rates of national industry employment over the year prior to each census, with the weights calculated as the share of a local labour market’s employment in each industry at the time of the census.

\textsuperscript{23} These four relationships are all as predicted. In particular, an inverse correlation between the renting proportion and the unemployment rate is consistent with Oswald’s (1997) explanation of unemployment.

\textsuperscript{24} That is, the correlation can only be due to the fact that the lag of $\ln u$ influences $\varepsilon$. 

14
of the combination of instruments, a positive and significant estimate of the unemployment elasticity of pay is obtained.

These results should be treated with some degree of caution. Blanchflower and Oswald’s conclusion that the wage curve is robust to controlling for simultaneity bias is possibly due to the high positive correlation that exists between observations in annual unemployment series such as the one they used. A consequence of this is that the one-year lagged unemployment rate the authors employ as an instrument closely resembles the endogenous contemporaneous unemployment rate and is unlikely to induce much deviation from the ordinary least squares estimates. In contrast, Table 4 reveals that the five-year lag of \( \ln u \) used here has a fairly weak correlation with the logarithm of current unemployment, which is not surprising given that the two values reflect different phases of the business cycle in each urban area. Since no annual unemployment series in New Zealand gives a sufficient degree of regional information, the instruments for \( \ln u \) used here are the best available.\(^{25}\) If a one-year lag on unemployment was obtainable, one might expect a similar conclusion to that of Blanchflower and Oswald.

If the results presented in the final column of Table 2 are assumed to be valid, they might suggest the rejection of a downward-sloping wage curve in favour of a short-run Harris-Todaro relationship. Alternatively, they may imply some form of misspecification of the earnings equation.\(^{26}\) However, the results are most indicative of the difficulty of finding appropriate instruments, rather than rejecting the wage curve.

As a final issue, it is informative to consider the extent to which the data used in this study have facilitated more accurate estimates of the wage curve than those of the earlier New Zealand work by Morrison and Poot (1999), which used aggregate census information. Obtaining disaggregated census data is a highly expensive undertaking; therefore it is useful to have some indication of the marginal benefits that are associated with using it to produce coefficient estimates. Lang and Gottschalk (1996) proposed a means to assess the efficiency loss that results from using grouped data to estimate the coefficients of variables that vary across groups, but not individuals.

\(^{25}\) The quarterly Household Labour Force Survey provides the official estimate of the unemployment rate in New Zealand, however this survey only began measuring regional unemployment in 1990 and, even then, only for 17 local government areas.

15
within groups, when individual data are not available on the dependent variable. They showed that use of disaggregated data only increases the precision of coefficient estimates for group-specific variables if there is correlation between the individual and group level regressors. In the current context, Lang and Gottschalk’s approach provides an estimate of the gain in efficiency attained by this study from using a “synthetic” microeconomic data set, compared with the aggregated data of Morrison and Poot.

The first step requires regressing the logarithm of the unemployment rate, $\ln u$, on all individual level regressors, namely the age, education, gender, ethnicity, employment status, occupation, urban area and year dummies. The $R^2$ from this regression is then compared with the $R^2$ from a similar regression, where the subgroup level variables are averaged so as to make them of the same level of aggregation as the dependent variable. This procedure was applied twice. Firstly, observations were aggregated within urban areas, years and gender, reducing the number of observations from 20,302 to 180. This resulted in the $R^2$ of interest rising from 0.871 to 0.937. Following Lang and Gottschalk, this implies that the standard error of $\hat{\beta}$ increases by 43% when switching to aggregate data. Next, $\ln u_T$ was used as the dependent variable, rather than $\ln u$, and observations were aggregated within urban areas and years only. This left 90 observations for the urban area level regression and an $R^2$ that rose from 0.952 to 0.983. It was determined that, when observations by urban area and year only are available, standard errors of $\hat{\beta}$ will be 69% higher than those presented in this study. Alternatively, it may be claimed that the use of disaggregated data has allowed estimates of the wage curve that are 41% more precise than those reported previously.

4. Conclusion

Having added to a diverse and rapidly expanding body of literature, what then has this paper contributed? Its primary aim was to investigate the role that local labour

---

26 Papps (2000) examines the specification of the wage curve in greater detail and finds that, for certain groups, a wage curve that is more consistent with theory is robust to controls for endogeneity.

27 The increase in the standard error of $\delta$ is equal to $\left(1 - R^2_n\right)^{1/7}$, where $R^2_n$ and $R^2_g$ denote the $R^2$ values of the group level and individual level regressions, respectively.
market conditions play in the determination of workers’ wages and, specifically, to test for evidence of a wage curve in New Zealand. As such, it represents the first study to use disaggregated labour market information, relating to the characteristics of particular subgroups of individuals.

Drawing on data from the 1986, 1991 and 1996 Censi of Population and Dwellings, a traditional Mincerian earnings equation was augmented to include the local unemployment rate as a regressor in various specifications. Consistent with the conclusion of Blanchflower and Oswald, evidence was found to favour the inclusion of the unemployment rate in logarithmic form. Using this specification, weighted least squares estimates provided evidence of a wage curve, with an unemployment elasticity of pay close to the internationally-reported value of –0.1. Support was also found for Harris and Todaro’s (1970) concept of a positive relationship between wages and unemployment in the long run, with the wage curve related to contemporaneous values of the two variables.

Evidence was found that the unemployment rate is endogenously determined with the wage level, indicating the presence of simultaneity bias. To overcome this, instrumental variables estimation was employed, using a set of instruments suggested by Blanchflower and Oswald. This procedure resulted in the unemployment elasticity of pay becoming positive, suggesting that the wage curve may not be robust to controls for the endogeneity of labour market conditions. However, some questions remain about the appropriateness of the instruments used.

This research has provided a new perspective on regional wage adjustment in New Zealand by its use of disaggregated census data. This was found to result in a substantial gain in efficiency over previous work by Morrison and Poot (1999) that employed aggregate data. To enable even more precise estimates of the wage curve, future studies should draw on individual level information on income and hours worked.

When considered alongside human capital variables, like education and job experience, labour market conditions may be regarded as a comparatively minor source of variation in earnings. Nonetheless, differences in the level of responsiveness across groups of workers or time periods can potentially provide a better insight into the determinants of labour market outcomes. To this end, Papps (2000) provided a first investigation and addressed some issues related to the mobility rates of different groups. Future work should extend this by including the specific unemployment and
migration rates for various subsets of the working age population that are obtainable from the census. These provide more accurate measures of unemployment probabilities and mobility levels, respectively, and may be more suitable variables to include in the earnings equation. This will increase the extent to which earnings and unemployment rates are endogenously related, thereby exacerbating a problem that was identified in this study. A final priority for research on the wage curve should therefore be to assess the most appropriate method of estimating the wage curve to allow for the presence of simultaneity bias.
### Appendix: Key to variables referred to

<table>
<thead>
<tr>
<th>Urban areas</th>
<th>Occupations</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Whangarei</td>
<td>1 Health professionals</td>
</tr>
<tr>
<td>2 Auckland</td>
<td>2 Teaching professionals</td>
</tr>
<tr>
<td>3 Hamilton</td>
<td>3 Office clerks</td>
</tr>
<tr>
<td>4 Tauranga</td>
<td>4 Salespersons, demonstrators and models</td>
</tr>
<tr>
<td>5 Rotorua</td>
<td>5 Personal and productive service workers</td>
</tr>
<tr>
<td>6 Gisborne</td>
<td>6 Skilled and semi-skilled production workers</td>
</tr>
<tr>
<td>7 Napier/Hastings</td>
<td>7 Construction workers</td>
</tr>
<tr>
<td>8 New Plymouth</td>
<td>8 Labourers and related elementary service workers</td>
</tr>
<tr>
<td>9 Wanganui</td>
<td></td>
</tr>
<tr>
<td>10 Palmerston North</td>
<td></td>
</tr>
<tr>
<td>11 Wellington</td>
<td></td>
</tr>
<tr>
<td>12 Nelson</td>
<td></td>
</tr>
<tr>
<td>13 Christchurch</td>
<td></td>
</tr>
<tr>
<td>14 Dunedin</td>
<td></td>
</tr>
<tr>
<td>15 Invercargill</td>
<td></td>
</tr>
<tr>
<td>16 Pukekohe</td>
<td></td>
</tr>
<tr>
<td>17 Tokoroa</td>
<td></td>
</tr>
<tr>
<td>18 Taupo</td>
<td></td>
</tr>
<tr>
<td>19 Whakatane</td>
<td></td>
</tr>
<tr>
<td>20 Hawera</td>
<td></td>
</tr>
<tr>
<td>21 Feilding</td>
<td></td>
</tr>
<tr>
<td>22 Levin</td>
<td></td>
</tr>
<tr>
<td>23 Kapiti</td>
<td></td>
</tr>
<tr>
<td>24 Masterton</td>
<td></td>
</tr>
<tr>
<td>25 Blenheim</td>
<td></td>
</tr>
<tr>
<td>26 Greymouth</td>
<td></td>
</tr>
<tr>
<td>27 Ashburton</td>
<td></td>
</tr>
<tr>
<td>28 Timaru</td>
<td></td>
</tr>
<tr>
<td>29 Oamaru</td>
<td></td>
</tr>
<tr>
<td>30 Gore</td>
<td></td>
</tr>
</tbody>
</table>
References


