Isolating the Scarring Effects Associated with the Economic Inactivity of Youth in New Zealand: Evidence from the Christchurch Health and Development Study

REPORT TO THE LABOUR MARKET POLICY GROUP

NEW ZEALAND DEPARTMENT OF LABOUR

3 March 2004

Tim Maloney
Research Economist
15 Staley Road
Parau, Auckland
t.maloney@auckland.ac.nz
Phone: 64-9-373-7599 ext.7597
Fax: 64-9-373-7427

Table of Contents

Executive Summary

This study estimates the potential scarring effects associated with early economic inactivity of young people in New Zealand. We initially define 'economic inactivity' as occurring when an individual is <u>not</u> enrolled in education or training, and <u>not</u> working in the labour market. Two basic variations on this definition are also considered: *(i)* excluding those living with a dependent child; and *(ii)* including those in part-time education, training or work as inactive. The term 'scarring' refers to the effects of this early economic inactivity on subsequent labour market outcomes. In this study, we examine the extent to which economic inactivity at ages 16, 18 and 21 influence the probability of being economically inactive at age 25, once detailed measures of personal and family background characteristics are held constant.

Data from the Christchurch Health and Development Study (CHDS) are used in this study. The CHDS provides multiple observations on possible inactivity from three interviews during the school-to-work transition period. It also contains extensive information on the family backgrounds, academic achievements and school characteristics for this cohort of subjects born in Canterbury hospitals in 1977. Both the range and quality of these independent variables are key features of this study. Their inclusion as explanatory variables in the regressions allows us to control for heterogeneity that might otherwise bias our estimates of these scarring effects.

Rates of economic inactivity in this sample increase from ages 16 to 21 as young people leave education, before declining slightly at age 25 as these subjects settle into their work careers. There is a clear 'path dependence' in the inactivity histories of youth. Nearly four-fifths of those in the sample were never economically inactive at the time of the four interviews. Those inactive at earlier ages are far more likely to be inactive at later ages.

Across all four definitions of economic inactivity, consistent evidence of a substantial scarring effect is found. Earlier inactivity is positively and significantly related to later inactivity, and this relationship is not eliminated by the inclusion of other explanatory variables. Larger scarring effects are estimated when we ignore both 'living with a dependent child' and 'part-time' education, training and work as forms of economic activity.

How important is it to have detailed information on personal and family backgrounds for isolating this scarring effect? Without any other control variables, economic inactivity at age 21 raises the probability of being inactive at age 25 by an estimated 25.3 percentage points for our base definition. Once the full set of controls are included in this regression, the estimated effect of inactivity at age 21 on the probability of inactivity at age 25 declines to 18.1 percentage points. Thus, the scarring effect is overestimated by nearly 40% when observed heterogeneity is not held constant. Yet, the magnitude of this effect is still substantial, more than doubling the actual proportion of individuals who are inactive at age 25.

Few of the background variables (e.g., IQ, classroom performance, conduct problems, peer associations and educational attainment, the parents' qualifications and work histories, and the family's structure, income, living standards and benefit history) are found to be individually significant in the regressions on economic inactivity at age

25. This may be due to the facts that inactivity measured at a single point in time may be a poor proxy for the permanent propensity to be inactive, and substantial collinearity exists among these explanatory variables.

Although the CHDS contains excellent data on personal and family backgrounds, this does not eliminate the possibility that (unobserved) heterogeneity may continue to overestimate the scarring effects. To address this potential omitted-variables problem, retrospective data on educational and work histories of young people between the interviews from ages 16 to 25 were added to these regressions. Under this specification, the only evidence of statistically significant and positive scarring effects is associated with definitions of inactivity that exclude those living with a dependent child from this group. Even though such auxiliary regressions may help narrow the range of estimates for potential scarring effects, further work in this area could better isolate the effects of early economic inactivity on long-term labour market outcomes for young people in New Zealand.

1. Introduction

The primary goal of this study is to measure the potential 'scarring effects' from the economic inactivity of young people in New Zealand. The hypothesis is that an early experience of economic inactivity will directly increase the probability that an individual will experience bouts of economic inactivity in the future.

One of the key issues in attempting to isolate this scarring effect is to acquire data on these experiences from the very beginning of this stochastic process. Ideally, we would like more or less continuous information on these labour force histories from the outset of this labour market transition (e.g., monthly or quarterly measures of inactivity). The problem is that these indicators of inactivity are likely to be poorly measured by retrospective data. Young people in the Christchurch Health and Development Study (CHDS) were interviewed at ages 16, 18, 21 and 25. These interviews provide both contemporaneous and retrospective data on their labour force activities and other life experiences. Our concern is that the reported labour force histories over the intervening periods may be incomplete and tainted by recall bias.

Individuals in the CHDS were asked whether they were working, enrolled in either education or training programmes, or unemployed at some point during each quarter between their 18^{th} and 21^{st} and 21^{st} and 25^{th} birthdays. This means that our only indicators of inactivity during this three-month period are unemployment or a complete lack of any employment, education or training over the quarter. We cannot distinguish between someone who worked for the entire three months and another person who worked for a single day and was out of the labour force for the balance of that quarter.

Scarring effects in the economics literature are more often associated with unemployment rather than economic inactivity. However, one of the limitations on unemployment in this context is that this state is only officially recognised for those who are actively in the labour force. Individuals who are both not working, and not actively seeking and available for employment are considered to be 'out of the labour force'. Unemployment incidence is unobserved for those outside of the labour force (e.g., those enrolled in education or training programmes). For these reason, much of the literature on scarring effects concentrates on prime-age males who are more likely to maintain a continuous attachment to the labour force over an extended observation period.

Of course, any association between unemployment spells over time for an individual could be indicative of either scarring effects or unobserved heterogeneity. The distinction between these sources for unemployment persistence is critical. The problem is that the stochastic process that generates this unemployment has been underway for some time before these prime-age males are observed. Early unemployment spells can capture both heterogeneity and state dependence. This is the so-called 'initial conditions problem'. The independent variable is correlated with a latent, person-specific effect in these regressions.

The potential advantage of the CHDS is that it follows the progress of a cohort of youth as they make the transition from education to work. In other words, we follow the progress of young people from the very beginning of the stochastic process that generates their eventual labour force histories.

Yet, several difficulties arise in accurately depicting the transition process between education and work for young people. This is partly due to the fact that these individuals are interviewed in the CHDS at intervals that increase steadily from 2 to 3 to 4 years between the ages of 16 and 25. More frequent interviews would make it easier to capture these initial forays into the labour market. But even more frequent data collection would not overcome the problem associated with young people moving back and forth between education and the labour market, and between being in and out of the labour force repeatedly over this transition period. The problem is that, unlike prime-age males, youth often display a fairly transitory attachment to the labour market.

For these reasons, we begin by redefining the research question. Instead of asking how bouts of unemployment might influence the subsequent probability of unemployment, we ask how 'economic inactivity' might affect the subsequent probability of this inactivity. Firstly, an absence of both human capital formation and work is probably closer to what most people consider to be the state that would have the most detrimental long-term effects on future employment opportunities. Secondly, these inactivity indicators are observed at every age, and our not conditional on labour force participation.

The remainder of this report is organised in the following way. The domestic and overseas economic literature on scarring effects are briefly summarised in Section 2. The econometric models used in this project are outlined in Section 3. The data and results from this regression analysis are discussed in Sections 4 and 5, respectively. General conclusions from this study are presented in Section 6.

2. A Brief Review of the Scarring Literature

One way to characterise economic inactivity is to emphasise its long-term 'scarring effects'. Early bouts of inactivity might reduce the rate of formal and informal human capital accumulation, damage self-esteem, and create poor personal and work habits. This inactivity may slow subsequent wage growth, and lead to more frequent bouts of inactivity in the future. The emphasis here is on the involuntary nature of economic inactivity. The inactivity of young people is potentially costly because human capital and work attitudes are particularly malleable early in one's life, and the amortisation period over which these poor returns will be received is particularly long.

Caspi et al. (1998) examined early failures in the school-to-work transition process for 954 youth by age 21 in the Dunedin Multidisciplinary Health and Development Study.[1] The dependent variable in their regression analysis was the proportion of months between the ages of 15 and 21 that these youth were unemployed (not employed but actively seeking work, and not a full-time student or homemaker). The

[1] There exists an earlier unpublished working paper prepared for the New Zealand Ministry of Youth Affairs (Silva and Caspi (1996)), which summarise the preliminary findings from this same study. Although the stated emphasis of this working paper was on the possible policy implications of this school-to-work transition, it provided no additional discussion along this dimension.

authors found that youth were more likely have been unemployed if they had poor reading skills and low IQ scores, and if they were raised in families with low income and a single parent. The risk of unemployment was also greater if these youth lacked an "attachment" to school, and demonstrated certain aspects of antisocial behaviour. No attempts were made in this study to link unemployment experiences at different ages.

Since some of these explanatory variables stretch back to early childhood, the authors conclude that factors that influence labour market outcomes by age 21 are already well established before youth begin their school-to-work transition process. They also note that many of these personal and family background characteristics have a significant impact on unemployment even after controlling for educational attainment. Thus, these factors have both direct and indirect effects (through education) on subsequent unemployment outcomes. Caspi et al. speculate that without the appropriate data on personal and family backgrounds it may be easy to overstate the detrimental effects of early unemployment experiences on poor labour market outcomes later in life.

Although Caspi et al. focused on the unemployment outcomes for all youth in this cohort over the six-year period between their 15th and 21st birthdays, the authors were actually looking at something closer to 'economic inactivity'. Nearly 20% of these youth were enrolled in tertiary study at age 21. As a result, their absence of unemployment by this age was largely due the fact that they were still enrolled in full-time study and had not yet started the labour market transition process.

Caspi et al. also claimed that several personal and family background measures had a significant impact on the unemployment incidence of youth after controlling for "human capital". However, their proxies for the educational attainment were surprisingly limited in this study. The human capital variables used were the occupational status of the parents, a dichotomous variable on whether or not the youth took the School Certificate exam at age 15, and the youth's reading achievement as measured by the Burt Word Reading Test. It is important to note that controls for academic achievement (e.g., grade point averages), or any subsequent formal qualifications or years of education were never included in this regression. This raises concerns about any conclusion that these background measures have a direct influence on the unemployment behaviour of youth. It may be that these effects operate primarily through both the quality and quantity of the education eventually obtained by these young people. This issue was never adequately addressed in their study.

Phelps (1972) suggested that policies that alleviate unemployment in the short run will tend to lower the unemployment rate in the long run. Reducing the incidence and duration of unemployment would prevent the erosion of human capital that can lead unemployment later in life. This is the essence of state dependence. Two otherwise identical individuals can have quite different unemployment histories later in life because they had different unemployment experiences early in life. Exposure to unemployment might alter preferences, prices or constraints that influence subsequent choices or outcomes in this area.

Many studies have shown that, at an individual level, the best predictor of unemployment in the future is unemployment in the past. Although this could be interpreted as evidence of state dependence, it is also consistent with a story based on heterogeneity in the population. Some individuals may be more or less predisposed to experience unemployment in every period. Unless this heterogeneity can be held constant in our regressions, past unemployment may appear to influence future unemployment through omitted variables. Unobserved heterogeneity can lead to "spurious" state dependence (Heckman 1981a).

The distinction between true state dependence (scarring), and unobserved heterogeneity is critical for many public policies. For example, if scarring effects exist, then temporary economic downturns can have substantial long-term consequences. As Phelps claimed, the natural unemployment rate in the future may depend on the effectiveness of the government in reducing the unemployment today. Cohorts of youth, who are allowed to experience unemployment in their first few years out of education, may see their employment and wage prospects diminished over their long working lives. If scarring effects are nonexistent, of course, the benefits of economic stabilisation, training and job search programmes will be reduced accordingly. True duration dependence would strengthen the case for all forms of intervention among the unemployed or inactive population.

The few, and somewhat dated, studies in the US on this issue have found no evidence of state dependence on the unemployment incidence of prime-age male household heads (Corcoran and Hill 1985) and young men (Heckman and Borjas 1980).[2] This contrasts with more recent studies from other countries that find substantial evidence of scarring behaviour. Narendranathan and Elias (1993) found clear and pervasive evidence of state dependence among young British men. Flaig at al. (1993) found similar results for prime-age German men. Recent studies by Arulampalam et al. (2000) and Gregg (2001) using British panel (British Household Panel Survey and National Child Development Survey, respectively) also uncovered strong evidence of scarring effects. Knights et al. (2002) using Australian panel data (Australian Longitudinal Survey) concluded that past unemployment significantly increases the prospects of future unemployment after controlling for both observable and unobservable individual differences.

Differences between the US and other countries in measured state dependence could arise for a number of reasons. Firstly, the US studies are based on data from the 1960's and 1970's, while the European and Australian studies come from data from the 1980's and 1990's. It is possible that state dependence in unemployment may be a relatively recent phenomenon. Secondly, the differences between the unemployment insurance and social welfare systems in the US and other countries may have something to do with these divergent findings. Stricter limitations on the duration of unemployment benefits and much tighter eligibility criteria for social welfare benefits in the US might eliminate any scarring effects that would otherwise exist. Finally,

[2] Gardecki and Neumark (1998) take a slightly different slant on this general issue of "churning" in early labour market experiences. Using data from the US National Longitudinal Survey of Youth, the authors find that labour market outcomes for adults in their late 20s and early 30s are largely unrelated to instability in early labour market experiences (e.g., holding frequent short-term, often dead-end jobs). As a result, they find little evidence to support policies that would bring more order to what often appears to be chaotic school-to-work transitions.

there a host of methodological and data issues that arise in this literature, and differences in many aspects of these studies could have lead to these conflicting findings. As Arulampalam et al. (2000) suggest *"... much more research is required in order to compare inter-country differences in state dependence in unemployment occurrence ... Such a research programme may shed light on why there are observed differences in the natural unemployment rate between the US and Europe."*

3. An Econometric Model of the Scarring Effects from Inactivity

A simple regression framework will be used for this analysis. Our plan is to use the detailed information in the CHDS on personal and family backgrounds to control for observed heterogeneity and to isolate any state dependence associated with several different measures of economic inactivity.

Consider the following regression specification.

$$I_{i25}^{*} = \alpha + Z_i'\beta + u_i \qquad (1)$$

The dependent variable is the latent propensity to be economically inactive at age 25. It is assumed to be a linear function of a vector Z_i of observable personal and family background characteristics for the individual and a disturbance term u_{i}. Of course, what we observe is the current economic inactivity of the individual. Either the person is inactive (I_{i25} =1) or active (I_{i25} =0) at the time of the 25-year interview. One of the key features of this analysis is the availability of a wide range of background factors in the CHDS that come from repeated surveys of parents, teachers and subjects from birth through age 25. The diversity and quality of these background measures should mitigate some of the unobserved heterogeneity that would otherwise exist in this estimation, and serve as good predictors for the probability of being economically inactive.

Past incidences of economic inactivity are then gradually added to this original model.

$$I_{i25}^{*} = \alpha + Z_i'\beta + \gamma_{21} I_{i21} + u_i \qquad (2)$$
$$I_{i25}^{*} = \alpha + Z_i'\beta + \gamma_{21} I_{i21} + \gamma_{18} I_{i18} + u_i \qquad (3)$$
$$I_{i25}^{*} = \alpha + Z_i'\beta + \gamma_{21} I_{i21} + \gamma_{18} I_{i18} + \gamma_{16} I_{i16} + u_i \qquad (4)$$

An initial indication of the potential scarring effects of economic inactivity by age 25 will come from the estimated coefficients on the binary variables associated with earlier incidences of inactivity for the individual. These dummy variables on economic inactivity are progressively added for ages 21, 18 and 16 to regressions (2) through (4).

Simple t and F tests will be used on the null hypotheses that the γ coefficients on these earlier measures of economic inactivity are individually or jointly equal to zero (indicating an absence of scarring effects), once other measurable background factors are held constant. In addition, the pattern in these γ coefficients will be of interest. We might expect that the effects of previous inactivity would weaken over time (i.e., $\gamma_{16} < \gamma_{18} < \gamma_{21}$). The inclusion of more recent indicators of inactivity may actually obscure the measurement of the influence of early indicators if the effects are cumulative in nature. In other words, the most recent measure of economic inactivity serves as a sufficient statistic for the history of inactivity for the individual up to that age.

The coefficients on the realisations of the lagged dependent variables are expected to capture the state dependence associated with past inactivity. In general, positive signs on the γ coefficients could arise from spurious state dependence for two reasons. Firstly, earlier episodes of inactivity could overlap two or more consecutive time periods. This is an artefact of data collection procedures where arbitrary time periods are used (e.g., adjacent calendar quarters or years). This is unlikely to be a serious problem in this situation, because of the long time intervals between observations on inactivity from the interviews at ages 25, 21 18 and 16.

Secondly, because of unobserved individual-specific factors, the disturbance term may be correlated with the lagged dependent variable. For example, an individual with little motivation or perseverance might have a higher permanent propensity to be inactive. This unobserved heterogeneity could not be separated from the state dependence in these specifications.

One of the keys in isolating the true scarring effects is the so-called "initial conditions problem" (e.g., see Heckman 1981a and 1981b, Flaig et al. 1993, Arulampalam et al. 2000 and Knights et al. 2002 for the background on this issue). This problem occurs because of the inclusion of a lagged dependent variable in this regression. The initial conditions problem arises because the start of the observation period in the data set used for this estimation does not coincide with the beginning of the stochastic process that generates the inactivity outcomes. In a sample of prime-age males, for example, the first observation on the labour force status may occur many years after these individuals first entered the labour market. Inactivity in the first observation in the sample may be due to either a history of inactivity or person-specific factors that are both unobserved.

The typical approach used in this literature is to simultaneously estimate the determinants of this initial observation and all subsequent observations on unemployment incidence. The estimation of this system of equations is computationally burdensome (see Arulampalam et al. 2000 and Knights et al. 2002 for two recent alternative approaches to this estimation).

The potential advantage of the CHDS data is that we follow a cohort of youth as they make the transition between education and the labour market. In other words, unlike almost other panel data sets, we follow these youth from essentially the beginning of the stochastic process that generates their eventual inactivity histories. This has, at least the potential, of offering a new perspective on this initial conditions problem. Essentially the first period of observation *is* the period in which this whole process is initiated. There can be, by definition, no state dependence in the first period.

However, this approach is heavily dependent on the 'quality' of the CHDS data available for these purposes. This data issue will requires some close attention in this project

One final concern over this estimation is the possibility that the estimates over scarring effects may be biased upward due to data collection procedures where the same inactivity spells can cross the time intervals used in measuring unemployment incidence. Corcoran and Hill (1985) suggest that persistence in individual unemployment histories can occur for three reasons: (1) unobserved heterogeneity, (2) state dependence (i.e., scarring) and (3) data collection procedures where the same unemployment spells can cross time intervals used in measuring unemployment incidence (e.g., adjacent calendar years). Unless we control for both (1) and (2), we will be unable to isolate these scarring effects.

A procedure used by Arulampalam et al. (2000) will be used here to address this issue of unemployment spells that span adjacent time periods. This essentially involves netting out the effects of these longer spells by using lagged dependent variables that are lagged more than one year. The longer the time interval between current and past unemployment incidence, the smaller is the effect of data collection procedures in biasing our estimates of state dependence.

4. Multiple Definitions of Economic Inactivity in the CHDS

The CHDS follows the subsequent progress of more than 1,200 children born in hospitals in the Canterbury region between April and August 1977. The unit of observation in this panel study is the child or 'subject'. Attempts have been made to follow these subjects through any change in the structure of the households in which they have resided. Interviews with the adults in these households were conducted annually from birth through age 16 of the CHDS child. Substantial interviews with these subjects took place at ages 16, 18, 21 and 25.

There are many ways of measuring economic inactivity from the data available in the CHDS. Four possible measures are given in Table 1. All share in common the notion that inactivity occurs when individuals are *not* enrolled in education or training programmes, and are *not* working at the time of the survey. The two variations on this definition are removing young adults with a dependent child from the inactive category, and adding those in part-time education, training and work to this inactive group.

Definition ***(B)*** is probably the most common and simplest measure of inactivity. Only 16 of the 813 young people in our sample (2.0%) were both out of school and not working at age 16.[3] Since one of these subjects was living with a dependent child at that age, 15 of the 813 individuals (1.8%) were inactive according to definition ***(A)*** (*not* in education, training or work, and *not* caring for a dependent child).

Nineteen youth (2.3%) were not in education, training or work *full-time* at age 16 (definition ***(D)***). The term 'Full-time' is defined here as being in education, training or work for 30 or more hours per week in these combined activities at the time of the survey. Again, one subject was living with a dependent child, so 18 young people (2.2%) met the criteria for economic inactivity under ***(C)*** (*not* in education, training or

[3] It should be noted that this truly represents the onset of inactivity among the subjects in this sample. All of the young people in the CHDS were enrolled in school (or were being home schooled) at age15.

work full-time, and *not* living with a dependent child). Thus, there is very little evidence of economic inactivity at age 16 when the vast majority of these subjects were still in secondary school (or some other type of education or training).

Table 1
Alternative Definitions of Economic Inactivity

	Proportion of Sample:			
At Age:	***(A)*** *Not in Education Training or Work & Not Living with Dependent Child*	***(B)*** *Not in Education Training or Work*	***(C)*** *Not in Full-Time Education Training or Work & Not Living with Dependent Child*	***(D)*** *Not in Full-Time Education Training or Work*
16	0.018	0.020	0.022	0.023
18	0.066	0.082	0.109	0.127
21	0.096	0.139	0.198	0.261
25	0.079	0.137	0.133	0.230

Notes: These data come from the Christchurch Health and Development Study. The sample is restricted to 813 subjects who remained in the CHDS from birth through age 25, and provide valid data on personal and family characteristics that will be used later in this study. The four measures of economic inactivity are defined around the time of the birthdays for these subjects at ages 16, 18, 21 and 25. Full-time education, training or work is defined as 30 or more hours in these combined activities at the time of the survey.

By age 18, somewhere between 6.6% and 12.7% of these same young people were economically inactive. At all ages, the lowest percentages are associated with definition ***(A)***, and the highest percentages are associated with definition ***(D)***. Inactivity clearly rises as youth begin to leave school.

All of these same measures on inactivity reach a peak at age 21, when they range between 9.6% and 26.1%. Living with a dependent child and being in full-time education, training or work both increase in relative importance in defining economic inactivity through age 21. In other words, the absolute gaps between definitions ***(B)*** and ***(A)*** and ***(D)*** and ***(B)*** increase steadily between ages 16 and 21.

Economic inactivity declines across all four definitions by age 25. This is despite the fact that educational enrolment levels decline monotonically through age 25.

Two general patterns emerge from these data. Firstly, rates of economic inactivity increase substantially between ages 16 and 21, before falling at least slightly by age 25. Secondly, restricting economic inactivity to individuals without a dependent child, and including part-time education, training or work as a form of inactivity (with or without a dependent child) become increasingly important considerations at least through age 21.

The results Table 2 represent a quick attempt to validate the importance of our economic inactivity measures at age 25. How closely do these four measures relate to the past unemployment experiences and most recent personal incomes for these young people? The first two rows of this table report the means of the estimated proportions of time that the subjects were unemployed between their 16th and 25th birthdays. This variable is constructed from retrospective reports on the number of months unemployed since the previous interview at ages 18, 21 and 25. Individuals who were economically inactive at age 25 had long-term unemployment propensities that were between 2 and 3 times higher than those who were active at that same age.

Table 2

'Validation' of the Importance of Economic Inactivity at Age 25

	Alternative Definitions			
At Age:	***(A)*** *Not in Education Training or Work & Not Living with Dependent Child*	***(B)*** *Not in Education Training or Work*	***(C)*** *Not in Full-Time Education Training or Work & Not Living with Dependent Child*	***(D)*** *Not in Full-Time Education Training or Work*
Mean Unemployment Propensity Ages 16 to 25 for those **Inactive** at Age 25	0.149	0.121	0.146	0.118
Mean Unemployment Propensity Ages 16 to 25 for those **Active** at Age 25	0.057	0.055	0.051	0.048
Mean Personal Income Ages 24 to 25 for those **Inactive** at Age 25	$20,542	$16,834	$18,226	$15,965
Mean Personal Income Ages 24 to 25 for those **Active** at Age 25	$30,538	$31,783	$31,525	$33,844

Notes: These data come from the Christchurch Health and Development Study. The sample is restricted to 813 subjects who remained in the CHDS from birth through age 25, and provide valid data on personal and family characteristics that will be used in this study. The four measures of economic inactivity are defined at the time of the birthdays for these subjects at age 25. Full-time education, training or work is defined as 30 or more hours in these combined activities at the time of the survey. Unemployment propensities are computed as the proportion of months between ages 16 and 25 that the subject was not working, but actively searching for work. Retrospective data on months unemployed since the previous interview are available at the interviews at ages 18, 21 and 25. Personal income comes from categorical information which is converted to dollar amounts using the midpoints of the income ranges over the closed intervals. The top, opened-ended income category is arbitrarily assigned a value of $120,000 for those receiving more than $100,000 in annual income. More than one in seven respondents lived overseas at age 25. Their personal incomes were converted to New Zealand dollars using the relevant Purchasing Power Parity (PPP) measure. These PPP rates equalize the purchasing power of incomes in other currencies, and get closer to a comparison of relative living standards between New Zealand and other countries. PPP figures come from the OECD (www.oecd.org).

The last two rows in Table 2 show substantial gaps between the personal incomes of economically active and inactive individuals over the preceding year. Incomes are between approximately 1.5 and 2 times higher for active subjects compared to inactive subjects at age 25.

The previous analysis is 'backward looking' showing the associations between unemployment and income histories and the most recently available measures of economic inactivity. However, we could also validate the importance of earlier observations on inactivity by 'looking forward' to see how these inactivity measures at an earlier age relate to subsequent important outcomes. Age 18 is chosen because of the small numbers of inactive youth at age 16. Table 3 shows that economic inactivity at age 18 is associated with are 3 to 4 times higher unemployment propensities over the following 7 years, and incomes that are approximately 50% lower at least six years after the interview at age 18 than the incomes of economically active individuals.

Table 3

'Validation' of the Importance of Economic Inactivity at Age 18

	Alternative Definitions			
At Age:	***(A)*** *Not in Education Training or Work & Not Living with Dependent Child*	***(B)*** *Not in Education Training or Work*	***(C)*** *Not in Full-Time Education Training or Work & Not Living with Dependent Child*	***(D)*** *Not in Full-Time Education Training or Work*
Mean Unemployment Propensity Ages 18 to 25 for those **Inactive** at Age 18	0.236	0.192	0.201	0.175
Mean Unemployment Propensity Ages 18 to 25 for those **Active** at Age 18	0.059	0.060	0.055	0.056
Mean Personal Income Ages 24 to 25 for those **Inactive** at Age 18	$21,351	$20,725	$21,257	$20,827
Mean Personal Income Ages 24 to 25 for those **Active** at Age 18	$30,336	$30,549	$30,785	$31,.036

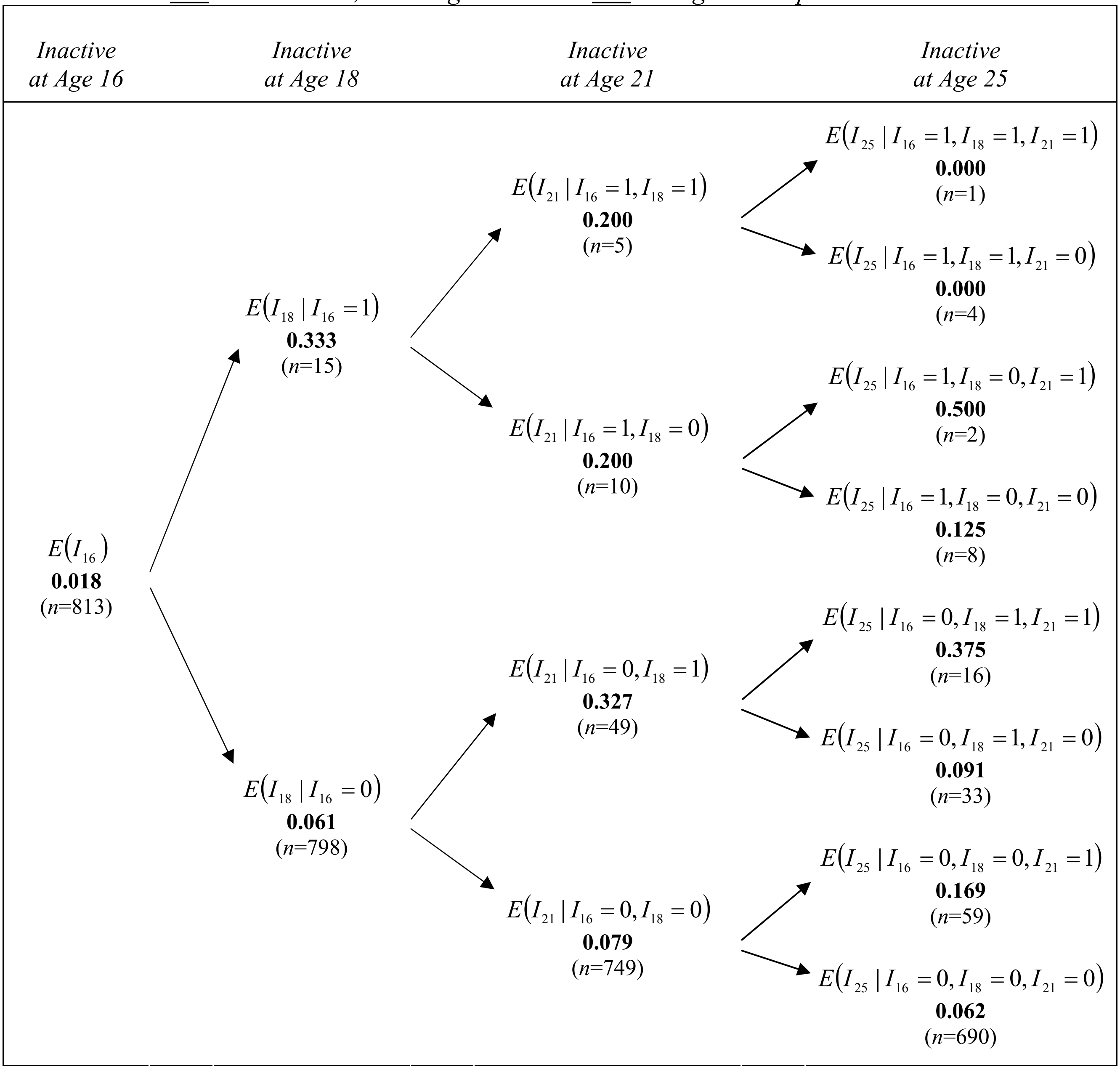
Notes: See the notes at the bottom of Table 2.

One of the advantages of the longitudinal data in the CHDS is displayed in Figure 1. It is typical in the scarring literature to look initially at the simple 'path dependence' of unemployment or economic inactivity histories. One way to do this is to compute the incidence of current inactivity conditional on inactivity at earlier ages. If both state dependence and heterogeneity aren't present, then these later proportions should be unrelated to their histories.

Figure 1

*Conditional Outcome Tree for Economic Inactivity using Definition **(A)***

<u>Not</u> in Education, Training or Work & <u>Not</u> Living with Dependent Child



Notes: The first column indicates the proportion of subjects who were economically inactive at the time of the interview at age 16. This group is then separated into those who did and did not experience inactivity at the time of three subsequent interviews. For example, the incidence of inactivity at age 18 varies substantially by measured inactivity two year earlier.

The proportions reported in this figure are the fractions of the 813 subjects who were economically inactive using definition ***(A)*** (i.e., *not* in education, training or work, and *not* living with a dependent child) at the time of each interview. Only 1.8% of subjects were inactive at age 16 (see Table 1). This figure appears in the first column of Figure 1. As a result of this low level of inactivity at age 16, the inactivity indicators at ages 18, 21 and 25 for these 15 individuals are displayed in the upper panel of this diagram. In other words, the vast majority of the young people in our sample were active at age 16 and their subsequent inactivity histories are shown in the bottom panel of Figure 1.

We already know from Table 1 that 6.6% of subjects were inactive (under this particular definition) at age 18. However, this incidence varies substantially by the earlier observation on inactivity. Only 6.1% of young people who were *not* inactive at age 16 were inactive at age 18. Exactly one-third of the very small number of inactive young people at age 16, were also inactive at age 18.

The remainder of this 'outcome tree' on economic inactivity at ages 21 and 25 are displayed in the remaining two columns of Figure 1. Clear path dependences exist in these statistics. At age 21, the probability of being economically inactive is 7.9% if the individual was active at ages 16 and 18. The probability of being economically inactive at the same age is 29.7% if the individual was inactive at either previous interview. At age 25, the probability of being economically inactive is 6.2% if the individual was active at ages 16, 18 and 21. The probability of being economically inactive is 17.1% at age 25 if the individual was inactive at the time of at least one of the three previous interviews.

Note that economic inactivity is concentrated among a relatively small subset of young people in our sample. Looking at the very bottom branch of this outcome tree in Figure 1, we see that 647 individuals (79.8% of our sample of 813) *never* experienced economic inactivity at the time of any of these four interviews. This means that observed inactivity is concentrated among approximately one-fifth of our sample. Yet, there is quite a bit of 'churning' in the incidence of inactivity. Nobody in this sample was economically inactive at every interview (the top branch in Figure 1). Only 9 individuals (1.1%) were inactive at three of the four interviews.

The estimated correlation coefficients on definition ***(A)*** at the time of these four interviews tell a similar story. They range from 0.147 to 0.168 for adjacent interviews (all significantly different from zero at a 1% level). They are 0.048 and 0.087 between the interviews at 16 and 21 and 18 and 21, respectively. Only the latter is statistically significant at a 1% level. The estimated correlation coefficient between inactivity at 16 and 25 is surprisingly close to zero (0.028), and not significantly different from zero at a 10% level.

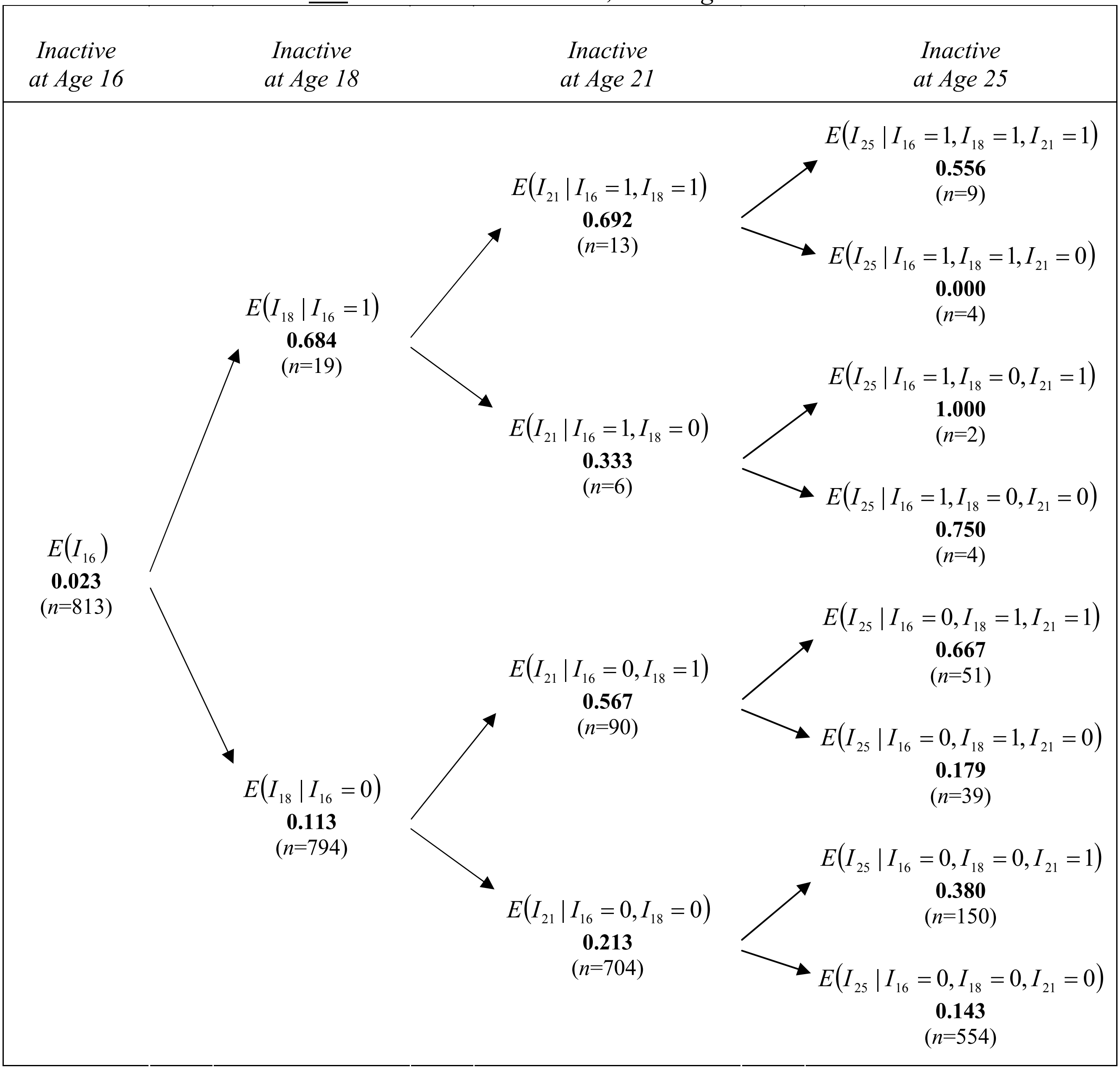
Figure 2 shows a similar tree diagram for definition ***(D)*** which results in the highest proportions of inactivity among our four definitions. Again, only 2.3% of the individuals in our sample were not in full-time education, training or work at age 16. At age 21, the probability of being economically inactive is 21.3% if the individual was active at ages 16 and 18. The probability of being inactive at the same age is 56.9% if the individual was inactive at either previous interview. At age 25, the probability of being inactive is 14.3% if the individual was active at ages 16, 18 and 21. The probability of being inactive is 41.7% at age 25 if the individual was inactive at one of the three previous interviews.

These results suggest that state dependence and/or heterogeneity in economic inactivity behaviour is relatively strong in the CHDS. This is consistent with the overseas scarring literature on unemployment.

Figure 2

*Conditional Outcome Tree for Economic Inactivity using Definition **(D)***

Not in Full-Time Education, Training or Work



Notes: The first column indicates the proportion of subjects who were economically inactive at the time of the interview at age 16. This group is then separated into those who did and did not experience inactivity at the time of three subsequent interviews. For example, the incidence of inactivity at age 18 varies substantially by measured inactivity two year earlier.

The estimated correlation coefficients on definition ***(D)*** are uniformly higher than those previously reported for definition ***(A)***. They range from 0.259 to 0.328 for adjacent interviews (all significantly different from zero at a 1% level). They are 0.112 and 0.196 between the interviews at 16 and 21 and 18 and 21, respectively. Both are significant at a 1% level. The estimated correlation coefficient between inactivity at 16 and 25 is 0.109, and significantly different from zero at a 1% level.

5. *Results from Regression Analysis on Economic Inactivity*

The results from the maximum likelihood probit estimation of equations (1) through (4) are shown in Table 4 for definition ***(A)*** on economic inactivity. The dependent variable is dichotomous. It equals one if the young person is *not* in education, training or work, and *not* living with a dependent child at age 25; zero otherwise. The estimated parameters are partial derivatives of the probability of economic inactivity for one-unit changes in the independent variables. The standard errors associated with these partial derivatives are in parentheses below these estimated derivatives.

When earlier indicators of economic inactivity are excluded from this regression (column 1 of Table 4), the only statistically significant factors that affect the probability of being inactive at age 25 are the mother's proportion of years working full-time and her mean depression score. The estimated effects of the other independent variables are not individually significantly different from zero at a 10% level. A full description of the explanatory variables used in this estimation can be found in Appendix A to this report.

Having a mother who worked full-time in all years between the ages 1 and 14 decreases the probability that her son or daughter will be economically inactive at age 25 by an average of 11.1 percentage points relative to a mother who never worked full-time over this period. This partial derivative is significantly different from zero at better than a 5% level. This estimated effect is approximately 1.4-times the size of the sample mean of 7.9% for this dependent variable.

An unexpected result is found on the Mean Depression Score for the mother. An increase of one standard deviation in maternal depressive symptoms (averaged over ages 6 through 13 for her CHDS child) is associated with a decrease in the probability that this subject will be economically inactive by 2.7 percentage points. This partial derivative is significant at better than a 5% level. To put this result in perspective, the estimated effect is approximately one-third of the sample mean.

One reason why the other independent variables have insignificant effects on the probability of economic inactivity is that these many personal and family background measures may be highly collinear. Multicollinearity increases standard errors and reduces our ability to reject the null hypotheses that individual coefficients are equal to zero. Yet, collinearity doesn't affect our R^2 statistics. Although an R^2 statistic cannot be computed under Probit estimation, we can produce a 'pseudo' R^2 statistic that approximates the explanatory power of the model.[4] Roughly 5.6% of the variation in the underlying probability of being economically inactive at age 25 can be explained by the variation in all of the detailed measures of personal and family backgrounds included in this regression. Thus, the vast majority of the variation in this economic behaviour cannot be explained by our extensive array of background variables in the CHDS.[5]

---

[4] See the note at the bottom of Table 4 for the formula, and associated reference, for the Estrella Pseudo R^2 Statistic.

[5] In auxiliary regressions not included in this report, the Grade Point Average from the individual's Sixth Form Certificate exams was added to the list of explanatory variables. The estimated coefficients on this variable were consistently insignificant in all regressions.

Table 4

Regression Results on the Probability of Being Economically Inactive at Age 25
*Defn. **(A)**: Not in Education, Training or Work & Not Living with Dependent Child*

Independent Variables	*Excluding Earlier Inactivity*	*Including Inactivity Age 21*	*Including Inactivity Ages 21 and 18*	*Including Inactivity Ages 21, 18 and 16*
Constant	0.052 (0.152)	0.020 (0.146)	0.012 (0.146)	0.005 (0.146)
Female	-0.005 (0.019)	-0.002 (0.018)	-0.003 (0.018)	-0.004 (0.018)
Maori or Pacific Islander	0.039 (0.032)	0.033 (0.031)	0.032 (0.030)	0.030 (0.030)
School Qualification Mother	0.036 (0.024)	0.035 (0.023)	0.034 (0.023)	0.033 (0.023)
Post-School Qualification Mother	0.053 (0.033)	0.050 (0.032)	0.048 (0.031)	0.047 (0.031)
School Qualification Father	0.017 (0.023)	0.011 (0.022)	0.012 (0.022)	0.011 (0.022)
Post-School Qualification Father	0.054 (0.033)	0.046 (0.031)	0.047 (0.031)	0.046 (0.031)
Number of Younger Siblings	0.003 (0.010)	0.003 (0.009)	0.002 (0.009)	0.003 (0.009)
Number of Older Siblings	0.010 (0.009)	0.010 (0.009)	0.009 (0.009)	0.009 (0.009)
Proportion Years Part-Time Work Mother	-0.018 (0.035)	-0.019 (0.033)	-0.017 (0.033)	-0.016 (0.033)
Proportion Years Full-Time Work Mother	-0.111^{**} (0.055)	-0.107^{**} (0.053)	-0.104^{**} (0.053)	-0.100^{*} (0.053)
Proportion Years Part-Time Work Father	-0.154 (0.331)	-0.146 (0.332)	-0.171 (0.336)	-0.182 (0.339)
Proportion Years Full-Time Work Father	0.060 (0.104)	0.077 (0.106)	0.068 (0.106)	0.068 (0.106)
Mean Depression Score Mother	-0.027^{**} (0.011)	-0.025^{**} (0.010)	-0.025^{**} (0.010)	-0.026^{**} (0.010)
Proportion Years in Two-Parent Family	-0.090 (0.097)	-0.104 (0.098)	-0.096 (0.098)	-0.093 (0.098)
Proportion Years Family on Benefit	0.016 (0.075)	0.009 (0.073)	0.007 (0.073)	0.007 (0.073)
Mean Real Family Income	0.017 (0.015)	0.017 (0.014)	0.016 (0.014)	0.016 (0.014)
Mean Family Living Standards	-0.025 (0.029)	-0.026 (0.028)	-0.025 (0.028)	-0.024 (0.028)
Mean IQ Test Score	-0.011 (0.013)	-0.006 (0.013)	-0.005 (0.013)	-0.005 (0.013)
Scholastic Ability Test Score	0.007 (0.014)	0.003 (0.014)	0.003 (0.014)	0.003 (0.013)
Mean Grade Point Average	-0.023 (0.016)	-0.019 (0.016)	-0.020 (0.016)	-0.020 (0.016)

Table 4 Continued

Mean Class Size	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)	-0.001 (0.002)
Proportion Years Private or Church School	0.003 (0.029)	0.001 (0.028)	0.001 (0.028)	-0.000 (0.028)
Association with Deviant Peers	-0.009 (0.010)	-0.010 (0.009)	-0.010 (0.009)	-0.011 (0.009)
Mean Conduct Problem Score	-0.005 (0.010)	-0.006 (0.010)	-0.008 (0.010)	-0.008 (0.010)
School Certificate	-0.055 (0.040)	-0.036 (0.036)	-0.036 (0.036)	-0.032 (0.036)
Sixth Form Certificate	0.002 (0.024)	0.000 (0.023)	0.003 (0.023)	0.004 (0.023)
Bursary	0.005 (0.028)	0.007 (0.027)	0.008 (0.027)	0.007 (0.027)
Post-School Qualification	0.014 (0.019)	0.019 (0.018)	0.020 (0.018)	0.020 (0.018)
University Degree	-0.018 (0.024)	-0.011 (0.025)	-0.010 (0.025)	-0.011 (0.025)
Economically Inactive at Age 21	---	0.119*** (0.048)	0.112** (0.047)	0.113** (0.047)
Economically Inactive at Age 18	---	---	0.033 (0.042)	0.034 (0.042)
Economically Inactive at Age 16	---	---	---	0.035 (0.076)
'Pseudo' R^2	0.056	0.070	0.071	0.072
N		813		

*** Significantly different from zero at 1% level, two-tailed test.
** Significantly different from zero at 5% level, two-tailed test.
* Significantly different from zero at 10% level, two-tailed test.

Notes: The dependent variable is a dummy variable that equals one if the subject was *not* in education, training or work and *not* living with a dependent child at age 25; zero otherwise. The dummy independent variables on economic inactivity at earlier ages correspond to this same definition. The estimated parameters reported in this table are related to the partial derivatives of the probability of being economically inactive with respect to each of the independent variables. The Pseudo R^2 statistic was developed by Estrella (1998, *Journal of Business and Economic Statistics*, 17). It is a function of the log-likelihood statistics with a constant (L_0) and with all independent variables (L):

$$\textit{Estrella Pseudo } R^2 \textit{ Statistic } = 1 - \left(\frac{L}{L_0}\right)^{-2L_0/N}$$

One set of surprising results from this regression are the consistently positive effects of the four dummy variables for parental qualifications on the probability of economic inactivity for their offspring. Although none of these coefficients are individually significant in Table 4, we can reject the null hypothesis that all four partial derivatives are simultaneously equal to zero at a 6.6% level.

The same regression model is re-estimated with the inclusion of a lagged dependent variable on economic inactivity from the previous interview at age 21. These results are reported in the second column of Table 4. The magnitudes of the estimated coefficients on the proportion of years of full-time work by the mother and her mean depression score both decline slightly in magnitude when this early measure of economic inactivity is included. However, the estimated derivatives on both variables continue to be statistically significant at better than a 5% level. None of the coefficients on the other background variables are individually statistically significant.

The estimated derivative on economic inactivity at age 21 is 0.119, and significantly different from zero at better than a 1% level. This says that, holding measured background factors constant, being economically inactive at age 21 raises the probability of being inactive at age 25 by an average of 11.9 percentage points. This is a relatively large effect considering the fact that the sample mean for inactivity by this definition is 7.9%. This finding provides some support for the early scarring effects associated with economic inactivity.

One way of assessing the importance of including regressors in this model to control for heterogeneity, is to drop all variables on personal and family backgrounds and compare the resulting coefficient estimate on inactivity at age 21 to the coefficient estimate on the same variable reported in Table 4. The estimated partial derivative on economic inactivity at age 21 increases from 0.119 to 0.154 and the Pseudo R^2 statistic falls from 0.070 to 0.021 when all other independent variables are excluded from this estimation. Both estimated coefficients on lagged inactivity are significantly different from zero at better than a 1% level. The interpretation is that, unless we control for the heterogeneity that is observable in this study, the estimated scarring effect would be biased upward by nearly 30%. Yet, it is just as important to note that this estimated scarring effects is not entirely eliminated by the inclusion of these other independent variables in this regression.

When the regression model is re-estimated with the inclusion of additional lagged dependent variables from age 18 (third column), and ages 18 and 16 (fourth column), these earlier effects are estimated to be positive, but not significantly different from zero. Among the measures of previous inactivity, only the measure of inactivity at age 21 is statistically significant. The estimated derivative on inactivity at age 21 declines only slightly from 0.119 to 0.112 when inactivity at age 18 is added to the equation. The estimated coefficient on this first lag remains largely unchanged (0.113) when inactivity at age 16 is also included in the estimation.

The inclusion of these indicators of previous economic inactivity raises the explanatory power of the model from 0.056 to at most 0.072. This comes almost entirely from the measure of inactivity at age 21.

Table 5 reports the results from the regressions when definition ***(B)*** on economic inactivity is used as our dependent variable. This variable equals one if the young person is *not* in education, training or work; zero otherwise. This definition may be closer to the traditional definition of economic inactivity, where having a dependent child in the household is *not* considered to be an equivalent ‘productive activity’.

Table 5

Regression Results on the Probability of Being Economically Inactive at Age 25
*Defn. **(B)**: Not in Education, Training or Work*

Independent Variables	*Excluding Earlier Inactivity*	*Including Inactivity Age 21*	*Including Inactivity Ages 21 and 18*	*Including Inactivity Ages 21, 18 and 16*
Constant	0.030 (0.203)	-0.053 (0.198)	-0.093 (0.200)	-0.100 (0.201)
Female	0.095*** (0.025)	0.079*** (0.025)	0.076*** (0.025)	0.075*** (0.025)
Maori or Pacific Islander	0.075* (0.040)	0.065* (0.039)	0.059 (0.039)	0.057 (0.039)
School Qualification Mother	0.028 (0.029)	0.025 (0.028)	0.024 (0.028)	0.023 (0.028)
Post-School Qualification Mother	0.016 (0.037)	0.012 (0.035)	0.009 (0.035)	0.008 (0.035)
School Qualification Father	0.016 (0.029)	0.008 (0.028)	0.007 (0.028)	0.007 (0.028)
Post-School Qualification Father	0.094** (0.044)	0.087** (0.043)	0.086** (0.043)	0.086** (0.043)
Number of Younger Siblings	0.008 (0.013)	0.009 (0.012)	0.010 (0.012)	0.010 (0.012)
Number of Older Siblings	0.022* (0.012)	0.023* (0.012)	0.022* (0.012)	0.022* (0.012)
Proportion Years Part-Time Work Mother	-0.054 (0.048)	-0.050 (0.047)	-0.042 (0.048)	-0.041 (0.048)
Proportion Years Full-Time Work Mother	-0.083 (0.071)	-0.085 (0.069)	-0.075 (0.069)	-0.072 (0.069)
Proportion Years Part-Time Work Father	-0.685 (0.472)	-0.689 (0.475)	-0.704 (0.480)	-0.719 (0.483)
Proportion Years Full-Time Work Father	0.007 (0.129)	0.032 (0.130)	0.028 (0.131)	0.025 (0.131)
Mean Depression Score Mother	-0.017 (0.012)	-0.022* (0.012)	-0.022* (0.012)	-0.022* (0.013)
Proportion Years in Two-Parent Family	-0.068 (0.116)	-0.074 (0.116)	-0.064 (0.117)	-0.060 (0.117)
Proportion Years Family on Benefit	0.009 (0.093)	0.010 (0.092)	0.009 (0.092)	0.008 (0.092)
Mean Real Family Income	0.011 (0.021)	0.010 (0.020)	0.008 (0.020)	0.008 (0.020)
Mean Family Living Standards	-0.031 (0.040)	-0.026 (0.039)	-0.021 (0.039)	-0.019 (0.039)
Mean IQ Test Score	-0.022 (0.018)	-0.016 (0.018)	-0.014 (0.018)	-0.014 (0.018)
Scholastic Ability Test Score	0.010 (0.020)	0.006 (0.019)	0.004 (0.019)	0.004 (0.019)
Mean Grade Point Average	-0.007 (0.022)	-0.004 (0.022)	-0.006 (0.022)	-0.006 (0.022)

Table 5 Continued

Mean Class Size	-0.001 (0.003)	-0.001 (0.003)	-0.001 (0.003)	-0.001 (0.003)
Proportion Years Private or Church School	-0.045 (0.043)	-0.050 (0.042)	-0.052 (0.042)	-0.053 (0.042)
Association with Deviant Peers	-0.007 (0.012)	-0.006 (0.012)	-0.009 (0.012)	-0.009 (0.012)
Mean Conduct Problem Score	-0.005 (0.013)	-0.011 (0.013)	-0.013 (0.013)	-0.013 (0.013)
School Certificate	-0.069 (0.046)	-0.039 (0.042)	-0.033 (0.042)	-0.030 (0.042)
Sixth Form Certificate	-0.052 (0.035)	-0.043 (0.034)	-0.034 (0.034)	-0.034 (0.034)
Bursary	0.035 (0.040)	0.043 (0.041)	0.045 (0.041)	0.045 (0.041)
Post-School Qualification	-0.014 (0.025)	-0.003 (0.025)	-0.000 (0.025)	-0.000 (0.025)
University Degree	-0.077*** (0.030)	-0.067** (0.030)	-0.067** (0.030)	-0.067** (0.030)
Economically Inactive at Age 21	---	0.181*** (0.049)	0.166*** (0.049)	0.168*** (0.050)
Economically Inactive at Age 18	---	---	0.078 (0.054)	0.076 (0.054)
Economically Inactive at Age 16	---	---	---	0.030 (0.083)
'Pseudo' R^2	0.094	0.120	0.123	0.124
N		813		

*** Significantly different from zero at 1% level, two-tailed test.
** Significantly different from zero at 5% level, two-tailed test.
* Significantly different from zero at 10% level, two-tailed test.

Notes: The dependent variable is a dummy variable that equals one if the subject was *not* in education, training or work at age 25; zero otherwise. The dummy independent variables on economic inactivity at earlier ages correspond to this same definition. The estimated parameters reported in this table are related to the partial derivatives of the probability of being economically inactive with respect to each of the independent variables. The Pseudo R^2 statistic was developed by Estrella (1998, *Journal of Business and Economic Statistics*, 17). It is a function of the log-likelihood statistics with a constant (L_0) and with all independent variables (L):

$$\textit{Estrella Pseudo } R^2 \textit{ Statistic } = 1 - \left(\frac{L}{L_0}\right)^{-2L_0/N}$$

As a direct result of this new definition, the estimated coefficients on being female are consistently positive and statistically significant at better than a 1% level in all four regressions. When earlier measures of inactivity are excluded from the equation, being female increases the probability of being economically inactive by 9.5 percentage points. This gender effect declines slightly with the inclusion of indicators on past inactivity. The comparison of the results in Tables 4 and 5 suggests that this gender effect on inactivity disappears when we redefine 'activity' to include the care of dependent children.

Maori or Pacific Islanders are more likely to be inactive by definition ***(B)***, even though this was not true of definition ***(A)***, but this effect is no longer statistically significant when more than one lagged variable on inactivity is included in the equation (columns 3 and 4 of Table 5).

A post-school qualification by the father is estimated to have a positive and statistically significant effect in all regressions reported in Table 5. This was not true of the narrower definition of inactivity used in the regressions of Table 4.

Although a mother's full-time work is estimated to reduce economic inactivity by definition ***(A)***, it has no measurable impact on inactivity by definition ***(B)***. This suggests that a mother's full-time work reduces the probability that her offspring will be inactive at age 25 by increasing the likelihood of living with a dependent child, and *not* increasing the likelihood of being in education, training or work.

The number of older siblings significantly increases the probability of being inactive by definition ***(B)***, but has no measurable impact on inactivity by definition ***(A)***. The effect of having older siblings seems to work entirely through decreasing the likelihood of being in education, training or work.

The effects associated with maternal depression decline slightly in magnitude and statistical significance in going from definition ***(A)*** to definition ***(B)***. This suggests that maternal depression reduces the probability of being inactive by increasing the likelihood of living with a dependent child, and *not* increasing the likelihood of being in education, training or work at age 25.

The estimated effects of having a university degree are negative and statistically significant at better than a 5% level in all four regressions reported in Table 5. Since the estimated derivatives on this same variable in Table 4 were insignificant, we can conclude that this effect is primarily associated with increasing the likelihood of being in education, training or work at age 25.

Similar results to those found in Table 4 are found for the inclusion of lagged indicators of economic inactivity in Table 5. The most recent measure of past inactivity at age 21 is positive and significantly different from zero in all regressions. The estimated effects of inactivity at ages 18 and 16 are positive, but insignificant. Overall, the pseudo R^2 statistics increase from 0.094, without any lagged measures of inactivity, to at most 0.124 when all three lags are included.

To assess the overall importance of controlling for heterogeneity, all variables on personal and family background characteristics were excluded from these regressions. When economic inactivity at age 21 is included as the only explanatory variable, the estimated coefficient on this variable is 0.253 and statistically significant at better than a 1% level. The estimated coefficient on this same variable falls to 0.181 when all other independent variables are included from this estimation (column 1 of Table 5). This suggests that, unless we control for observable heterogeneity, the estimated scarring effect would be biased upward by nearly 40%.

Table 6

Regression Results on the Probability of Being Economically Inactive at Age 25

*Defn. **(C)**: Not in Full-Time Education, Training or Work, & Not Living with a Dependent Child*

Independent Variables	*Excluding Earlier Inactivity*	*Including Inactivity Age 21*	*Including Inactivity Ages 21 and 18*	*Including Inactivity Ages 21, 18 and 16*
Constant	-0.033 (0.210)	-0.090 (0.205)	-0.102 (0.205)	-0.119 (0.205)
Female	-0.024 (0.025)	-0.016 (0.025)	-0.019 (0.025)	-0.021 (0.025)
Maori or Pacific Islander	0.027 (0.038)	0.028 (0.037)	0.026 (0.037)	0.023 (0.036)
School Qualification Mother	0.008 (0.029)	0.007 (0.028)	0.005 (0.028)	0.003 (0.028)
Post-School Qualification Mother	0.025 (0.036)	0.016 (0.035)	0.013 (0.034)	0.009 (0.034)
School Qualification Father	0.002 (0.029)	-0.005 (0.028)	-0.005 (0.028)	-0.006 (0.028)
Post-School Qualification Father	0.088** (0.043)	0.084** (0.042)	0.084** (0.042)	0.082** (0.042)
Number of Younger Siblings	0.010 (0.013)	0.008 (0.013)	0.008 (0.013)	0.009 (0.013)
Number of Older Siblings	0.011 (0.013)	0.007 (0.013)	0.007 (0.013)	0.007 (0.012)
Proportion Years Part-Time Work Mother	0.036 (0.048)	0.044 (0.047)	0.047 (0.047)	0.051 (0.047)
Proportion Years Full-Time Work Mother	-0.071 (0.072)	-0.064 (0.071)	-0.062 (0.071)	-0.051 (0.071)
Proportion Years Part-Time Work Father	-0.323 (0.433)	-0.385 (0.435)	-0.408 (0.438)	-0.448 (0.442)
Proportion Years Full-Time Work Father	0.072 (0.139)	0.047 (0.139)	0.042 (0.139)	0.032 (0.138)
Mean Depression Score Mother	-0.027* (0.013)	-0.020 (0.013)	-0.020 (0.013)	-0.021 (0.013)
Proportion Years in Two-Parent Family	-0.096 (0.132)	-0.088 (0.132)	-0.080 (0.132)	-0.065 (0.133)
Proportion Years Family on Benefit	0.120 (0.095)	0.074 (0.094)	0.074 (0.094)	0.071 (0.094)
Mean Real Family Income	0.027 (0.021)	0.025 (0.021)	0.024 (0.021)	0.023 (0.020)
Mean Family Living Standards	-0.007 (0.039)	-0.002 (0.038)	-0.002 (0.038)	-0.001 (0.038)
Mean IQ Test Score	-0.020 (0.018)	-0.011 (0.018)	-0.011 (0.018)	-0.011 (0.018)
Scholastic Ability Test Score	0.019 (0.020)	0.009 (0.019)	0.008 (0.019)	0.008 (0.019)
Mean Grade Point Average	-0.037* (0.022)	-0.030 (0.022)	-0.030 (0.022)	-0.032 (0.022)

Table 6 Continued

Mean Class Size	-0.001 (0.002)	-0.001 (0.003)	-0.001 (0.003)	-0.001 (0.003)
Proportion Years Private or Church School	-0.016 (0.041)	-0.021 (0.040)	-0.019 (0.040)	-0.024 (0.040)
Association with Deviant Peers	-0.006 (0.013)	-0.008 (0.012)	-0.009 (0.012)	-0.011 (0.012)
Mean Conduct Problem Score	0.001 (0.013)	-0.003 (0.013)	-0.004 (0.013)	-0.006 (0.013)
School Certificate	-0.096* (0.054)	-0.077 (0.051)	-0.075 (0.051)	-0.065 (0.050)
Sixth Form Certificate	0.056* (0.030)	0.056** (0.028)	0.059** (0.028)	0.062** (0.028)
Bursary	0.028 (0.039)	0.025 (0.038)	0.029 (0.038)	0.028 (0.038)
Post-School Qualification	0.020 (0.025)	0.016 (0.025)	0.018 (0.025)	0.018 (0.025)
University Degree	-0.052* (0.032)	-0.054* (0.030)	-0.053* (0.030)	-0.053* (0.030)
Economically Inactive at Age 21	---	0.174*** (0.038)	0.170*** (0.038)	0.169*** (0.038)
Economically Inactive at Age 18	---	---	0.036 (0.043)	0.036 (0.043)
Economically Inactive at Age 16	---	---	---	0.137 (0.117)
'Pseudo' R^2	0.061	0.097	0.098	0.101
N		813		

*** Significantly different from zero at 1% level, two-tailed test.
** Significantly different from zero at 5% level, two-tailed test.
* Significantly different from zero at 10% level, two-tailed test.

Notes: The dependent variable is a dummy variable that equals one if the subject was *not* in full-time education, training or work (i.e., 30 hours or more in the combined activities) and *not* living with a dependent child at age 25; zero otherwise. The dummy independent variables on economic inactivity at earlier ages correspond to this same definition. The estimated parameters reported in this table are related to the partial derivatives of the probability of being economically inactive with respect to each of the independent variables. The Pseudo R^2 statistic was developed by Estrella (1998, *Journal of Business and Economic Statistics*, 17). It is a function of the log-likelihood statistics with a constant (L_0) and with all independent variables (L):

$$\textit{Estrella Pseudo } R^2 \textit{ Statistic } = 1 - \left(\frac{L}{L_0} \right)^{-2L_0/N}$$

Table 6 reports the regression results when definition ***(C)*** for economic inactivity is used as our dependent variable. This dependent variable equals one if the young person is *not* in full-time education, training or work, and *not* living with a dependent child at age 25; zero otherwise. Full-time is defined as 30 or more usual hours in these combined activities.

Relative to the results from Table 4, school and post-school qualifications for the subject increase in importance when inactivity includes part-time education, training and work. When no measures of past inactivity are included in the equation, receiving a School Certificate and a university degree both significantly reduce the probability of being inactive at age 25.

Somewhat surprisingly in Table 5, a Sixth Form Certificate is found to increase the probability of inactivity. The significant estimated effects associated with a Sixth Form Certificate and university degree persist when past indicators of inactivity are included in the estimation.

Evidence of scarring effects associated with inactivity at age 21 can be found in Table 6. The estimated partial derivatives on this variable are found to decline slightly from 0.174 to 0.169 when earlier measures of inactivity are included in this analysis. None of the estimated coefficients on inactivity at ages 18 or 16 are statistically significant.

To assess the importance of controlling for heterogeneity, all variables on personal and family backgrounds were excluded from these regressions. When economic inactivity at age 21 is included as the only explanatory variable, the estimated coefficient on this variable is 0.198 and statistically significant at better than a 1% level. The estimated coefficient on this same variable falls to 0.174 when all other independent variables are included from this estimation (column 1 of Table 6). This suggests that, unless we control for observable heterogeneity, the estimated scarring effect would be biased upward by nearly 14%.

Table 7 reports the regression results when definition ***(D)*** for economic inactivity is used as our dependent variable. This dependent variable equals one if the subject is *not* in full-time education, training or work at age 25; zero otherwise. Full-time is again defined as 30 or more usual hours in these combined activities.

Relative to the results from Table 5, school and post-school qualifications for the subject appear to play a larger role in being active full-time, rather than just having some attachment to education, training or work. Obtaining a School Certificate or university degree significantly lowers the probability of being out of full-time education, training or work at age 25. Regression results show that a Sixth Form Certificate increased the probability of economic inactivity under definition ***(C)***, but not definition ***(D)***. These results suggest that a Sixth Form Certificate increases the probability of being inactive by reducing the likelihood of living with a dependent child, and *not* reducing the likelihood of being in full-time education, training or work.

Evidence of scarring effects associated with inactivity at age 21 is also apparent in Table 7. The estimated partial derivatives on this variable are found to decline very slightly from 0.265 to 0.257 when earlier measures of inactivity are included in this analysis. None of the estimated coefficients on inactivity at ages 18 or 16 are statistically significant.

Table 7
Regression Results on the Probability of Being Economically Inactive at Age 25
Defn. ***(D)****: Not in Full-Time Education, Training or Work*

Independent Variables	*Excluding Earlier Inactivity*	*Including Inactivity Age 21*	*Including Inactivity Ages 21 and 18*	*Including Inactivity Ages 21, 18 and 16*
Constant	0.067 (0.269)	-0.080 (0.270)	-0.106 (0.271)	-0.125 (0.272)
Female	0.170*** (0.032)	0.145*** (0.032)	0.139*** (0.033)	0.139*** (0.033)
Maori or Pacific Islander	0.071 (0.047)	0.064 (0.047)	0.059 (0.047)	0.056 (0.047)
School Qualification Mother	0.000 (0.036)	-0.001 (0.036)	-0.003 (0.036)	-0.005 (0.036)
Post-School Qualification Mother	-0.013 (0.044)	-0.028 (0.043)	-0.033 (0.043)	-0.035 (0.043)
School Qualification Father	-0.014 (0.036)	-0.023 (0.036)	-0.023 (0.036)	-0.023 (0.036)
Post-School Qualification Father	0.115** (0.052)	0.128** (0.053)	0.126** (0.053)	0.125** (0.053)
Number of Younger Siblings	0.030* (0.017)	0.029* (0.017)	0.030* (0.017)	0.031* (0.017)
Number of Older Siblings	0.034** (0.016)	0.031* (0.016)	0.030* (0.016)	0.031* (0.016)
Proportion Years Part-Time Work Mother	0.020 (0.064)	0.039 (0.064)	0.044 (0.064)	0.047 (0.064)
Proportion Years Full-Time Work Mother	0.029 (0.089)	0.047 (0.091)	0.055 (0.091)	0.064 (0.092)
Proportion Years Part-Time Work Father	-0.459 (0.518)	-0.573 (0.536)	-0.607 (0.542)	-0.642 (0.544)
Proportion Years Full-Time Work Father	0.079 (0.175)	0.051 (0.178)	0.055 (0.179)	0.040 (0.179)
Mean Depression Score Mother	-0.020 (0.016)	-0.019 (0.016)	-0.019 (0.016)	-0.019 (0.016)
Proportion Years in Two-Parent Family	-0.160 (0.161)	-0.121 (0.164)	-0.117 (0.165)	-0.098 (0.166)
Proportion Years Family on Benefit	0.153 (0.119)	0.088 (0.121)	0.088 (0.122)	0.084 (0.122)
Mean Real Family Income	0.033 (0.028)	0.026 (0.028)	0.025 (0.028)	0.024 (0.028)
Mean Family Living Standards	-0.025 (0.053)	-0.021 (0.052)	-0.011 (0.052)	-0.006 (0.053)
Mean IQ Test Score	-0.018 (0.025)	-0.010 (0.025)	-0.010 (0.025)	-0.010 (0.025)
Scholastic Ability Test Score	0.028 (0.026)	0.012 (0.027)	0.011 (0.027)	0.011 (0.027)
Mean Grade Point Average	-0.036 (0.029)	-0.029 (0.029)	-0.030 (0.029)	-0.031 (0.029)

Table 7 Continued

Mean Class Size	-0.003 (0.003)	-0.003 (0.003)	-0.003 (0.003)	-0.003 (0.003)
Proportion Years Private or Church School	-0.010 (0.053)	-0.022 (0.053)	-0.022 (0.053)	-0.027 (0.054)
Association with Deviant Peers	0.008 (0.015)	0.002 (0.015)	0.000 (0.016)	-0.001 (0.016)
Mean Conduct Problem Score	0.005 (0.017)	-0.005 (0.018)	-0.006 (0.018)	-0.008 (0.018)
School Certificate	-0.130** (0.059)	-0.107* (0.058)	-0.103* (0.058)	-0.095 (0.059)
Sixth Form Certificate	0.003 (0.042)	0.031 (0.040)	0.038 (0.040)	0.041 (0.040)
Bursary	0.037 (0.050)	0.038 (0.050)	0.044 (0.051)	0.043 (0.051)
Post-School Qualification	-0.009 (0.033)	-0.011 (0.033)	-0.007 (0.033)	-0.007 (0.033)
University Degree	-0.149*** (0.038)	-0.147*** (0.037)	-0.145*** (0.038)	-0.146*** (0.037)
Economically Inactive at Age 21	---	0.265*** (0.040)	0.257*** (0.041)	0.258*** (0.041)
Economically Inactive at Age 18	---	---	0.059 (0.052)	0.052 (0.052)
Economically Inactive at Age 16	---	---	---	0.103 (0.118)
'Pseudo' R^2	0.129	0.190	0.192	0.193
N		813		

*** Significantly different from zero at 1% level, two-tailed test.
** Significantly different from zero at 5% level, two-tailed test.
* Significantly different from zero at 10% level, two-tailed test.

Notes: The dependent variable is a dummy variable that equals one if the subject was *not* in full-time education, training or work (i.e., 30 hours or more in the combined activities) at age 25; zero otherwise. The dummy independent variables on economic inactivity at earlier ages correspond to this same definition. The estimated parameters reported in this table are related to the partial derivatives of the probability of being economically inactive with respect to each of the independent variables. The Pseudo R^2 statistic was developed by Estrella (1998, *Journal of Business and Economic Statistics*, 17). It is a function of the log-likelihood statistics with a constant (L_0) and with all independent variables (L):

$$\textit{Estrella Pseudo } R^2 \textit{ Statistic } = 1 - \left(\frac{L}{L_0}\right)^{-2L_0/N}$$

To assess the importance of controlling for observable heterogeneity, all variables on personal and family backgrounds were excluded from these regressions. When economic inactivity at age 21 is included as the only explanatory variable, the estimated coefficient on this variable is 0.314 and statistically significant at better than a 1% level. The estimated coefficient on this same variable falls to 0.265 when all other independent variables are included from this estimation (column 1 of Table 7). This suggests that, unless we control for observable heterogeneity, the estimated scarring effect would be biased upward by nearly 19%.

One of the unexpected findings in the regression results reported in Tables 4 through 7 are the consistently positive effects of parental qualification on the probability of the subject being economically inactive at age 25. For example, we've already noted that the four coefficients on the dummy variables related to the school and post-school qualifications of the parents are simultaneously different from zero in the regressions results displayed in Table 4. In Tables 5 through 7, a post-school qualification for the father has a consistently positive estimated effect that is significantly different from zero at better than a 5% level.

There are several explanations for this unexpected result. Firstly, it could be related to systematic measurement error on economic inactivity. Children of highly educated parents might be more likely to be in education themselves at age 25, but this may be coded as a form of inactivity if the subjects are temporarily away from education. For example, they may be on summer holiday at the time of the interview. Although we can't completely rule out this possibility, it is unlikely to stem from observations taking place during the summer break. Subjects in the CHDS are interviewed between the months of April and September.

Secondly, parental qualifications may have indirect effects through a number of other measures of personal and family backgrounds that are already included in these regressions. Examples would be performance on cognitive achievement tests, classroom performances and school and post-school qualifications. If we eliminated these mediating factors from the regression, the hypothesis is that parental qualifications would capture the overall negative effects on inactivity that we would expect.

It is possible to test this second hypothesis. In regressions not displayed in tables in this report, these possible mediating variables were excluded from all previously estimated regressions. The only background variables, other than indicators of earlier inactivity, included in these specifications were gender, ethnicity, numbers of younger and older siblings and the four dummy variables on parental qualifications.

These regressions show no evidence of consistently positive effects from parental qualifications on the probability of economic inactivity for their children at age 25. The *F* tests indicate that the null hypothesis that the four coefficients on parental qualifications are equal to zero cannot be rejected at even a 10% level in every regression. These results show that a post-school qualification for the father continues to have positive and significant effects on economic inactivity using definitions ***(C)*** and ***(D)***. However, these positive effects are largely offset by negative effects coming from the mother's post-school qualification and the father's school qualification.

Although the results of these auxiliary regressions support the claim that the earlier positive effects associated with parental qualifications were ignoring some of the negative effects from this education operating through mediating factors, they do *not* support the conclusion that parental qualifications (both directly and indirectly) reduce the probability of economic inactivity for their son or daughter by age 25.

A final conjecture for why parental education doesn't significantly reduce the probability of economic inactivity for their offspring could be the subject of future empirical work on this topic. It may be that children from highly educated households have 'higher expectations' in both their education and work careers. They are more likely to take time off from education, and may engage in more protracted job search as young adults. Failure to achieve these higher expectations may result in measured inactivity around age 25. Thus, although having parents with higher education levels may provide better opportunities for young people, it may also result in higher expectations for both educational and labour market achievement.

One final specification issue is considered here. Up to this point, little use has been made of the retrospective data in the CHDS from the surveys at ages 18, 21 and 25. Respondents were asked detailed questions at these interviews about their educational and work histories since the previous interview. We conjecture that the measured effects attributed to indicators of earlier inactivity may be more closely related to these educational and work histories. More specifically, we want to include continuous measures of the effective, full-time years of both education and work experiences for these subjects in all previous regressions. These are estimates of the amount of time that the individual spent in both education and work, using both retrospective and contemporaneous data from the interviews at ages 16, 18, 21 and 25. Specific definitions for these new independent variables are included in Appendix A.

To conserve on space, only the regression results related to these new explanatory variables and the indicators of past inactivity are displayed in Tables 4B through 7B in Appendix B. Table 4B shows that the amount of time spent in full-time education has no measurable impact on the probability of being economically inactive using definition ***(A)***. Remember that dummy variables on the school and post-school qualifications obtained are already included in this regression, along with all previous covariates. One year of full-time equivalent work since age 16 is estimated to reduce the probability of being inactive at age 25 by 2.2 percentage points when indicators of past economic inactivity are excluded. This estimated derivative declines slightly when these lagged dependent variables are included in the regression. All of these estimated effects are significant at better than a 1% level.

When variables on the education and work experience are added to these regressions, the estimated scarring effects decline substantially in both magnitude and statistical significance. The estimated partial derivatives on inactivity at age 21 vary between 0.063 and 0.065 in Table 4B. They are statistically significant at a 10% level. The inclusion of these additional regressors has substantially lowered our estimates of this state dependence. The estimated coefficients on this same variable using the same definition of inactivity in Table 4 were nearly twice as large as those reported in Table 4B (ranging from 0.112 to 0.119).

Although the addition of these more continuous variables on education and work experience does call into question our earlier estimates of large scarring effects, their inclusion in these regressions may also result in a downward bias on our estimates of the state dependence associated with inactivity. Economic inactivity from the periods between previous interviews would be expected to influence current inactivity. Yet, this inactivity would be partly reflected in these measures of cumulative education and work experience. It is unclear what proportion of the effects from these experience variables should be attributed to scarring.

It could be argued that the inclusion of the education and work experience variables produces 'lower bound' estimates for these scarring effects. The estimated partial derivatives on inactivity indicators from earlier interviews provide a minimum estimate of scarring once we hold constant personal and family backgrounds and comprehensive education and work histories. Previous estimates that did not control for education and work experiences might be thought of as 'upper bound' estimates for scarring. Under this reasoning, inactivity at age 21 is estimated to directly increase the probability of inactivity at age 25 by somewhere between 6.3 and 11.9 percentage points.

Tables 5B through 7B in Appendix B provide similar lower bound estimates for scarring effects using definitions ***(B)*** through ***(D)*** on economic inactivity. Once education and work experience are held constant, the estimated coefficients on all lagged dependent variables are insignificantly different from zero under definition ***(B)*** (see Table 5B). At a minimum, there may be no measurable scarring using this definition of inactivity.

Yet, under this same specification, the estimated derivatives on inactivity at age 21 are all positive and statistically significant under definition ***(C)*** in Table 6B. Combining these results with those from Table 6, we estimate that inactivity at age 21 directly increases the probability of inactivity at age 25 by somewhere between 10.2 and 17.4 percentage points.

Once education and work experience are held constant, the estimated coefficients on inactivity at age 21 are insignificantly different from zero with definition ***(D)*** (see Table 7B). One set of unusual results in these regressions is that inactivity at age 18 has negative and significant effects on inactivity at age 25.

6. Conclusions

The purpose of this study has been to provide some of the first estimates for the potential scarring effects associated with the early economic inactivity among young people in New Zealand. Once other relevant factors are held constant, is there any statistical evidence that current inactivity directly increase the probability of future inactivity?

Longitudinal data are required for this estimation because we need to link the economic inactivity of the same individuals over a sufficiently long period of time. These panel data should provide extensive information on personal and family backgrounds to control for heterogeneity in the population. Any remaining unobserved heterogeneity may bias our estimates of these scarring effects.

The Christchurch Health and Development Study (CHDS) satisfies the requirement of providing extensive data on the backgrounds and experiences of young people as they make their transitions between education and the labour market.

Several definitions of 'economic inactivity' have been used in this project. They all share the common view that inactivity occurs primarily when an individual is not in education, training or work at the time of the survey. Two variations on this basic definition include removing adults living with children from the inactive group, and adding those in part-time education, training and work to this category.

This study concentrates on the binary indicators of economic inactivity at the time of the surveys at ages 16, 18, 21 and 25 for the subjects in the CHDS. Rates of inactivity are found to increase substantially between ages 16 and 21, before declining at least slightly by age 25. Approximately 7.9% of the young people in our sample were *not* in education, training or work and *not* living with a dependent child at age 25. Yet, around 23.0% of these same individuals were *not* in full-time education, training or work at the same age.

There is clear evidence of path dependence in the inactivity histories of the young people in our sample. Indications of inactivity at an earlier age are associated with higher probabilities of inactivity at a later age. This suggests that heterogeneity or state dependence (or both) are behind this observed behaviour. Nearly four-fifths of our sample did not experience economic inactivity at the time of the interviews at ages 16, 18, 21 and 25. Yet, there is evidence of considerable churning in inactivity with no single individual being inactive at the time of all four surveys.

Although the CHDS provides a wealth of detailed information on personal and family background factors, very few of these variables are found to be individually significant in the regressions on economic inactivity at age 25. This lack of statistical significance may be attributed to the fact that a binary variable of inactivity at a single point in time is a fairly 'noisey' measure of any permanent or long-run propensity to be economically inactive. Only tiny fractions of the variations in the probabilities of being inactive can be explained by variations in background factors such as the subject's gender, ethnicity, IQ, classroom performance, conduct problems, peer associations and educational attainment, the parents' qualifications and work histories, and the family's structure, income, living standards and benefit history.

Holding these background measures constant, however, does reduce the magnitudes of our estimated scarring effects. If this extensive information is excluded form our regressions, the estimated derivatives on economic inactivity at age 21 increase by around 20 to 40%. Yet, there is strong and consistent evidence in our basic regressions across all four definitions that being inactive at the survey at age 21 is positively and significantly related to the probability of being inactive at age 25. This measured effect is largely unaffected by the inclusion of other lagged dependent variables from ages 18 and 16. These indicators of much earlier inactivity consistently have no measurable impact on this outcome at age 25.

The estimated scarring effects are found to be larger in magnitude when we ignore both 'living with a dependent child' and 'part-time' education, training and work as forms of economic activity. State dependence is more closely associated with being out of education, training and work, especially 30 or more hours per week in these combined activities. Different definitions of inactivity can produce different outcomes.

One unusual finding from these initial regressions was that parental qualifications did not have the anticipated negative effects on the economic inactivity of their children at age 25. In fact, generally these effects were found to be positive. Auxiliary regressions from this study show that this positive effect is largely due to the inclusion of mediating factors that pick up the indirect effects of parental education in reducing youth inactivity.

Yet, there is no statistical evidence in this study that parental qualifications directly and indirectly lower the rate of economic inactivity for offspring at age 25. Our conjecture is that although higher education levels for parents provide some additional opportunities for their children that should reduce this inactivity, they might also raise expectations of children in both education and the labour market that result in offsetting increases in inactivity at age 25. This claim is not tested empirically tested in this study, but could be the focus of some follow-up work on this topic.

One key issue with the results reported thus far that is common to similar studies in this area is that unobserved heterogeneity may be biasing upward the estimated scarring effects. Although the CHDS may have better data on personal and family backgrounds than past studies, this does not completely eliminate the possibility of this form of omitted-variable bias.

One way to address this potential problem is to add variables to these regressions that utilise retrospective data on educational and work histories of young people from the periods between the four interviews from ages 16 to 25. The idea is that the incidence of economic inactivity at age 25 may be more closely associated with these more continuous measures of effective full-time years of education and work experiences. Although it is difficult to know where to draw the line between economic inactivity (included to capture state dependence) and education and work experience (included to capture heterogeneity), it might be best to think of this as a way of producing lower bound estimates of scarring effects.

Once both background factors and education and work histories are included in these regressions, the only evidence of statistically significant and positive scarring effects comes from definitions of inactivity that include living with dependent children as a form of productive activity. Thus, although they may help narrow the range of estimates for these scarring effects, more work is needed on this approach to produce better estimates of the long-term consequences of early bouts of economic inactivity among young people in New Zealand.

References

Arulampalam, Wiji. (2001) "Is Unemployment Really Scarring? Effects of Unemployment Experiences on Wages." *The Economic Journal*, 111: F585-F606.

Arulampalam, Wiji, Paul Gregg and Mary Gregory. (2001) "Unemployment Scarring." *The Economic Journal*, 111: F577-F584.

Arulampalam, Wiji, Alison L. Booth and Mark P. Taylor. (2000) "Unemployment Persistence." *Oxford EconomicPapers*, 52: 24-50.

Chamberlain, Gary. (1984) "Panel Data." S. Griliches and M. Intriligator (eds.), in *Handbook of Econometrics*, North-Holland, Amsterdam: 1247-1318.

Corcoran, Mary and Martha S. Hill. (1985) "Reoccurrence of Unemployment Among Young Men." *Journal of Human Resources*, 20(2): 165-183.

Ellwood, David. (1982) "Teenage Unemployment: Temporary Scars or Permanent Blemishes." R.B. Freeman and D.A. Wise (eds.), in *The Youth Labor Market Problem: Its Nature, Causes and Consequences*, University of Chicago Press, Chicago, IL.

Flaig, Gerhard, Georg Licht and Viktor Steiner. (1993) "Testing for State Dependence Effects in a Dynamic Model of Male Unemployment Behaviour." H. Brunzel, P. Jensen and N. Westergard-Neilsen (eds.), in *Panel Data and Labour Market Dynamics*, North-Holland, Amsterdam: 189-213.

Gardecki, Rosella and David Neumark. (1998) "Order from Chaos? The Effects of Early Labor Market Experiences on Adult Labor Market Outcomes." *Industrial and Labor Relations Review*, 51(2): 299-322.

Gregg, Paul. (2001) "The Impact of Youth Unemployment on Adult Unemployment in the NCDS." *The Economic Journal*, 111: F626-F653.

Heckman, James J. (1981a) "Heterogeneity and State Dependence." S. Rosen (ed.), in *Studies in Labor Markets*, University of Chicago Press, Chicago, IL: 91-139.

Heckman, James J. (1981b) "The Incidental Parameters Problem and the Problem of Initial Conditions in Estimating a Discrete Time-Discrete Data Stochastic Process." C.F. Manski and D, McFadden (eds.), in *Structural Analysis of Discrete Data with Econometric Applications*, MIT Press, Cambridge, MA: 179-195.

Heckman, James J. and George J. Borjas. (1980) "Does Unemployment Cause Future Unemployment? Definitions, Questions and Answers from a Continuous Time Model of Heterogeneity and State Dependence." *Economica*, 47: 247-283.

Horwood, L.J. and D.M. Fergusson. (1986) "Neuroticism, Depression and Life Events: A Structural Equation Model." *Social Psychiatry*, 21: 63-71.

Knights, Stephen, Mark N. Harris and Joanne Loudes. (2002) "Dynamic Relationships in the Australian Labour Market: Heterogeneity and State Dependence." *The Economic Record*, 78(242): 284-298.

Narendranathan, W. and P. Elias. (1993) "Influences of Past History on the Incidence of Youth Unemployment: Empirical Findings for the UK." *Oxford Bulletin of Economics and Statistics*, 55: 161-185.

Phelps, Edmund S. (1972) *Inflation Policy and Unemployment*. Norton, New York.

Appendix A

Definitions of Independent Variables used in Regression Analysis

Variable Names	***Descriptions***
Female	Binary variable equal to one if the subject is female; zero if male.
Maori or Pacific Islander	Binary variable equal to one if the subject is Maori or Pacific Islander; zero for non-Maori and non-Pacific Islander.
School Qualification Mother	Binary variable equal to one if the highest educational qualification of the mother at the birth of the child is a school qualification; zero otherwise.
Post-School Qualification Mother	Binary variable equal to one if the highest educational qualification of the mother at the birth of the child is a post-school qualification; zero otherwise.
School Qualification Father	Binary variable equal to one if the highest educational qualification of the father at the birth of the child is a school qualification; zero otherwise.
Post-School Qualification Father	Binary variable equal to one if the highest educational qualification of the father at the birth of the child is a post-school qualification; zero otherwise.
Number of Younger Siblings	Maximum number of siblings younger than the subject living in the household in which the subject resided through age 15.
Number of Older Siblings	Maximum number of siblings older than the subject living in the household in which the subject resided through age 15.
Proportion Years Part-Time Work Mother	Proportion of annual interviews from ages 1 through 14 of the subject in which the mother or female custodial adult worked fewer than 30 hours per week.
Proportion Years Full-Time Work Mother	Proportion of annual interviews from ages 1 through 14 of the subject in which the mother or female custodial adult worked 30 or more hours per week.
Proportion Years Part-Time Work Father	Proportion of annual interviews from ages 1 through 14 of the subject in which the father or male custodial adult worked fewer than 30 hours per week.
Proportion Years Full-Time Work Father	Proportion of annual interviews from ages 1 through 14 of the subject in which the father or male custodial adult worked 30 or more hours per week.

Appendix A Continued

Mean Depression Score Mother	Mean of maternal depression score from ages 6 through 13 of the subject. In each of the eight years, mothers were questioned about their depressive symptoms over the month preceding the interview. Questions were based on the Levine-Pilowsky Depression Inventory. The scale originally ranged from 0 to 37 with high scores indicating symptoms of depression for the mother. This variable is standardized to have a zero mean and unit variance within our sample. See Horwood and Fergusson (1977) for background on this measure.
Proportion Years in Two-Parent Family	Proportion of years between the ages 1 and 14 in which the subject lived in a two-parented family.
Proportion of Years on Benefit	Proportion of years between ages 1 and 14 of the subject in which either parent was in receipt of social welfare benefits. These benefits came primarily from the Unemployment and Domestic Purposes Benefit.
Mean Real Family Income	This is the average real family between ages 1 and 14 of the child. The Consumer Price Index is used to inflate estimated family income from both labour and nonlabour sources at the time of each survey to constant 1996 dollars. This variable is standardized to have a zero mean and unit variance within our sample.
Mean Family Living Standards	This is the average of subjective impressions of CHDS interviewers over the family's standard of living at the time of the interviews when the subject was between 1 and 12 years. A five-point scaled is used, where 5 indicates a family that is "… obviously affluent or well to do," and 1 indicates a family that is "… obviously poor or very poor."
Mean IQ Test Score	Mean score on the Revised Weschler Intelligence Scale for Children administered by the CHDS when these children were aged 8 and 9 years. This variable is standardized to have a zero mean and unit variance within our sample.
Scholastic Ability Test Score	Test score on the Test of Scholastic Abilities (TOSCA) administered by the CHDS when the child was age 13. This test is designed to measure the extent to which the subject has the aptitudes necessary for success in high school. This variable is standardized to have a zero mean and unit variance within our sample.
Mean Grade Point Average	Mean Grade Point Average (GPA) of the subject between ages 7 and 12. Classroom teachers of these children were asked to rate their performance in the areas of reading, writing, spelling and mathematics over these 6 surveys. A five-point scale was used ranging from 1 for very poor to 5 for very good. The number reported here is the mean of these 4 variables across the 6 years.
Mean Class Size	Mean class size of subject between ages 7 and 17.
Proportion Years Private of Church School	Proportion of years between ages of 7 and 16 that the subject was in a private or church school.

Appendix A Continued

Association with Deviant Peers	At age 15 subjects were asked about their association with peers displaying various forms of deviant behaviour. A checklist was created with a minimum score (zero) indicating no deviant behaviour, and a maximum score (10) indicating substantial deviant behaviour among peers. This variable is standardized to have a zero mean and unit variance within our sample.
Mean Conduct Problem Score	Mean score on conduct problem surveys of both parents and teachers at ages 7, 9, 11 and 13. These conduct problems could include disruptive, oppositional, destructive and aggressive behaviour, as well as lying, stealing and cheating. This variable is standardized to have a zero mean and unit variance within our sample.
School Certificate	Binary variable equal to one if the subject obtained a School Certificate qualification by age 25; zero otherwise.
Sixth Form Certificate	Binary variable equal to one if the subject obtained a Sixth Form Certificate qualification by age 25; zero otherwise.
Bursary	Binary variable equal to one if the subject obtained a Bursary qualification by age 25; zero otherwise.
Post-School Qualification	Binary variable equal to one if the subject obtained a post-school qualification below a university degree by age 25; zero otherwise. This includes certificates, diplomas and other post-school qualifications.
University Degree	Binary variable equal to one if the subject obtained a university undergraduate or post-graduate degree by age 25; zero otherwise.
Years of Full-Time Education	This variable is estimated for this project from information taken at the time of the interviews at ages 16, 18, 21 and 25. Retrospective and contemporaneous data are used to estimate the amount of full-time education acquired by age 25. The resulting variable is measured years of effective, full-time education. One year of part-time education is assumed to equal one-half year of full-time education.
Years of Full-Time Work Experience	This variable is estimated for this project from information taken at the time of the interviews at ages 16, 18, 21 and 25. Retrospective and contemporaneous data are used to estimate the amount of full-time work experience by age 25. The resulting variable is measured years of effective, full-time work. One year of part-time work is assumed to equal one-half year of full-time work.

Appendix B

Table 4B

Regression Results on the Probability of Being Economically Inactive at Age 25
Defn. ***(A)****: Not in Education, Training or Work & Not Living with Dependent Child*
Including Explanatory Variables on Years of Education and Work

Independent Variables	*Excluding Earlier Inactivity*	*Including Inactivity Age 21*	*Including Inactivity Ages 21 and 18*	*Including Inactivity Ages 21, 18 and 16*
Years of Full-Time Education	-0.008 (0.005)	-0.006 (0.005)	-0.007 (0.006)	-0.006 (0.006)
Years of Full-Time Work Experience	-0.022*** (0.004)	-0.020*** (0.004)	-0.020*** (0.004)	-0.020*** (0.004)
Economically Inactive at Age 21	---	0.063* (0.038)	0.064* (0.039)	0.065* (0.039)
Economically Inactive at Age 18	---	---	-0.008 (0.025)	-0.006 (0.026)
Economically Inactive at Age 16	---	---	---	0.034 (0.070)
'Pseudo' R^2	0.095	0.101	0.101	0.101
		813		

*** Significantly different from zero at 1% level, two-tailed test.
** Significantly different from zero at 5% level, two-tailed test.
* Significantly different from zero at 10% level, two-tailed test.

Notes: The dependent variable is a dummy variable that equals one if the subject was *not* in education, training or work and *not* living with a dependent child at age 25; zero otherwise. The dummy independent variables on economic inactivity at earlier ages correspond to this same definition. The estimated parameters reported in this table are related to the partial derivatives of the probability of being economically inactive with respect to each of the independent variables. Other variables on personal and family backgrounds were included in these regressions (see Table 4 for comparison), but these results are not included in this table. The Pseudo R^2 statistic was developed by Estrella (1998, *Journal of Business and Economic Statistics*, 17). It is a function of the log-likelihood statistics with a constant (L_0) and with all independent variables (L):

$$\textit{Estrella Pseudo } R^2 \textit{ Statistic } = 1 - \left(\frac{L}{L_0}\right)^{-2L_0/N}$$

Table 5B

Regression Results on the Probability of Being Economically Inactive at Age 25
Defn. ***(B)***: *Not in Education, Training or Work*
Including Explanatory Variables on Years of Education and Work

Independent Variables	*Excluding Earlier Inactivity*	*Including Inactivity Age 21*	*Including Inactivity Ages 21 and 18*	*Including Inactivity Ages 21, 18 and 16*
Years of Full-Time Education	-0.032*** (0.008)	-0.031*** (0.008)	-0.032*** (0.008)	-0.032*** (0.008)
Years of Full-Time Work Experience	-0.050*** (0.006)	-0.049*** (0.007)	-0.051*** (0.007)	-0.051*** (0.007)
Economically Inactive at Age 21	---	0.010 (0.033)	0.011 (0.033)	0.012 (0.033)
Economically Inactive at Age 18	---	---	-0.028 (0.029)	-0.028 (0.029)
Economically Inactive at Age 16	---	---	---	0.014 (0.071)
'Pseudo' R^2	0.205	0.205	0.206	0.206
		813		

*** Significantly different from zero at 1% level, two-tailed test.
** Significantly different from zero at 5% level, two-tailed test.
* Significantly different from zero at 10% level, two-tailed test.

Notes: The dependent variable is a dummy variable that equals one if the subject was *not* in education, training or work at age 25; zero otherwise. The dummy independent variables on economic inactivity at earlier ages correspond to this same definition. The estimated parameters reported in this table are related to the partial derivatives of the probability of being economically inactive with respect to each of the independent variables. Other variables on personal and family backgrounds were included in these regressions (see Table 5 for comparison), but these results are not included in this table. The Pseudo R^2 statistic was developed by Estrella (1998, *Journal of Business and Economic Statistics*, 17). It is a function of the log-likelihood statistics with a constant (L_0) and with all independent variables (L):

$$\textit{Estrella Pseudo } R^2 \textit{ Statistic } = 1 - \left(\frac{L}{L_0}\right)^{-2L_0/N}$$

Table 6B

Regression Results on the Probability of Being Economically Inactive at Age 25
Defn. ***(C)****: Not in Full-Time Education, Training or Work, & Not Living with a Dependent Child*
Including Explanatory Variables on Years of Education and Work

Independent Variables	*Excluding Earlier Inactivity*	*Including Inactivity Age 21*	*Including Inactivity Ages 21 and 18*	*Including Inactivity Ages 21, 18 and 16*
Years of Full-Time Education	-0.008 (0.008)	-0.006 (0.008)	-0.007 (0.008)	-0.006 (0.008)
Years of Full-Time Work Experience	-0.045*** (0.006)	-0.040*** (0.006)	-0.040*** (0.006)	-0.040*** (0.006)
Economically Inactive at Age 21	---	0.102*** (0.034)	0.104*** (0.034)	0.104*** (0.034)
Economically Inactive at Age 18	---	---	-0.021 (0.031)	-0.018 (0.031)
Economically Inactive at Age 16	---	---	---	0.105 (0.110)
'Pseudo' R^2	0.139	0.155	0.156	0.158
	813			

*** Significantly different from zero at 1% level, two-tailed test.
** Significantly different from zero at 5% level, two-tailed test.
* Significantly different from zero at 10% level, two-tailed test.

Notes: The dependent variable is a dummy variable that equals one if the subject was *not* in full-time education, training or work (i.e., 30 hours or more in the combined activities) and *not* living with a dependent child at age 25; zero otherwise. The dummy independent variables on economic inactivity at earlier ages correspond to this same definition. The estimated parameters reported in this table are related to the partial derivatives of the probability of being economically inactive with respect to each of the independent variables. Other variables on personal and family backgrounds were included in these regressions (see Table 6 for comparison), but these results are not included in this table. The Pseudo R^2 statistic was developed by Estrella (1998, *Journal of Business and Economic Statistics*, 17). It is a function of the log-likelihood statistics with a constant (L_0) and with all independent variables (L):

$$\textit{Estrella Pseudo } R^2 \textit{ Statistic } = 1-\left(\frac{L}{L_0}\right)^{-2L_0/N}$$

Table 7B

Regression Results on the Probability of Being Economically Inactive at Age 25
Defn. ***(D)****: Not in Full-Time Education, Training or Work*
Including Explanatory Variables on Years of Education and Work

Independent Variables	*Excluding Earlier Inactivity*	*Including Inactivity Age 21*	*Including Inactivity Ages 21 and 18*	*Including Inactivity Ages 21, 18 and 16*
Years of Full-Time Education	-0.063*** (0.012)	-0.061*** (0.012)	-0.067*** (0.012)	-0.067*** (0.012)
Years of Full-Time Work Experience	-0.117*** (0.010)	-0.113*** (0.011)	-0.120*** (0.011)	-0.120*** (0.011)
Economically Inactive at Age 21	---	0.037 (0.038)	0.038 (0.038)	0.039 (0.038)
Economically Inactive at Age 18	---	---	-0.108*** (0.031)	-0.109*** (0.031)
Economically Inactive at Age 16	---	---	---	0.060 (0.120)
'Pseudo' R^2	0.360	0.362	0.370	0.371
	813			

*** Significantly different from zero at 1% level, two-tailed test.
** Significantly different from zero at 5% level, two-tailed test.
* Significantly different from zero at 10% level, two-tailed test.

Notes: The dependent variable is a dummy variable that equals one if the subject was *not* in full-time education, training or work (i.e., 30 hours or more in the combined activities); zero otherwise. The dummy independent variables on economic inactivity at earlier ages correspond to this same definition. The estimated parameters reported in this table are related to the partial derivatives of the probability of being economically inactive with respect to each of the independent variables. Other variables on personal and family backgrounds were included in these regressions (see Table 7 for comparison), but these results are not included in this table. The Pseudo R^2 statistic was developed by Estrella (1998, *Journal of Business and Economic Statistics*, 17). It is a function of the log-likelihood statistics with a constant (L_0) and with all independent variables (L):

$$\textit{Estrella Pseudo } R^2 \textit{ Statistic } = 1 - \left(\frac{L}{L_0}\right)^{-2L_0/N}$$