This is an author produced version of a paper published in Oxford Bulletin of Economics and Statistics. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the published paper:
Matz Dahlberg and Magnus Gustavsson
"Inequality and Crime: Separating the Effects of Permanent and Transitory Income"
URL: http://dx.doi.org/10.1111/j.1468-0084.2007.00492.x

Access to the published version may require subscription.
Published with permission from: Oxford Bulletin of Economics and Statistics
The definitive version is available at www.blackwell-synergy.com,
Inequality and Crime: Separating the Effects of Permanent and Transitory Income*

MATZ DAHLBERG† and MAGNUS GUSTAVSSON‡

† Department of Economics, Uppsala University, The Institute for Labour Market Policy Evaluation (IFAU), and CESifo
(e-mail: matz.dahlberg@nek.uu.se)

‡ Department of Economics, Uppsala University, Box 513, 751 20 Uppsala, Sweden
(e-mail: magnus.gustavsson@nek.uu.se)

Abstract
Earlier studies on income inequality and crime have typically used total income or total earnings. However, it is quite likely that it is changes in permanent rather than in transitory income that affects crime rates. The purpose of this paper is therefore to disentangle the two effects by, first, estimating region-specific inequality in permanent and transitory income and, second, estimating crime equations with the two separate income components as explanatory variables. The results indicate that it is important to separate the two effects; while an increase in the inequality in permanent income yields a positive and significant effect on total crimes and three different property crimes, an increase in the inequality in transitory income has no significant effect. Using a traditional, aggregate, measure of income yields insignificant effects on crime.

JEL classification: C33, D31, J39, K40

Keywords: Crime, Earnings dynamics, Inequality

I. Introduction
Earlier studies on income inequality and crime have typically used inequality in total income or total earnings as explanatory variables. However, income can be considered as consisting of two parts, one permanent and one transitory, and it is quite likely that it is changes in the permanent part rather than in the transitory part that affects crime rates, as the two have different ramifications for the duration of inequality. An increase in the dispersion of

* We have received useful comments and suggestions from Patrik Hesselius, Matthew Lindquist, Eva Mörk, an anonymous referee, and from seminar participants at the 2nd Nordic Summer Institute in Labor Economics, FIEF, SOFI and Uppsala University. Magnus Gustavsson gratefully acknowledges financial support from the Jan Wallander and Tom Hedelius Foundation.
permanent income leads to greater income inequality in both the short- and the long-term. An increase in the dispersion of transitory income, on the other hand, only creates short-term inequality.

The purpose of this paper is to disentangle the effects from inequality in permanent income from the effects from inequality in transitory income on crime. This is done in two steps. In the first step, we estimate region-specific inequality in permanent and transitory income using a very rich dataset on Swedish individuals. While several previous studies have decomposed year-to-year changes in inequality into its permanent and transitory components, this is the first time that these estimations have been carried out on a regional level. In the second step, we estimate crime equations with the two estimated income components as explanatory variables. This is, as far as we know, the first time that this separation of income has been used in the literature estimating the effects of income inequality on crime.

The results indicate that it is important to separate the two effects; while an increase in the inequality in permanent income yields a positive and significant effect on property crime, an increase in the inequality in transitory income has no significant effect. When using a traditional, composite, measure of income, we get insignificant effects on crime.

The rest of the paper is organized as follows. Section 2 presents some theoretical considerations on the relationship between income inequality and crime, and section 3 presents the income data and estimates permanent and transitory income for each county in Sweden. Section 4 discusses the econometric specification of the crime equations and the data to be used when estimating them. Section 5 presents the results, and, finally, section 6 concludes.

II. Income inequality and crime: Theoretical considerations

Why should there be an association between crime rates and income inequality? It turns out that both economic and sociological theory has linked income inequality to criminal activity. While economists have suggested that inequality may capture differential returns to criminal activity and, thereby, have an association with crime rates, sociologists have hypothesized that inequality and social welfare in general may work through other channels; inequality may be associated with lack of social capital, lack of upward mobility, or social disorganization, all of which may cause higher levels of crime.2

2 The discussion of these mechanisms follows Demombynes and Özler (2005).
Ever since Becker’s seminal work in the late 1960s (Becker, 1968), economists have suggested that economic incentives for crimes are higher in areas with greater income inequality (see e.g. Ehrlich, 1973; Chiu and Madden, 1998; Bourguignon, 2001; and Chisholm and Choe, 2005). In his work, Becker (1968) proposes an occupational choice model in which the incentives for individuals to commit crime are determined by the differential returns from legitimate and illegitimate pursuits.\(^3\)

However, based on economic theory, the sign of the effect of income inequality on crime is ambiguous. This is because income inequality may also be associated with the level of protection from crime. Private crime protection measures may include guard dogs, bars on windows, electric fences, and alarm systems with armed security response. Chiu and Madden (1998) provide a model that allows for richer neighborhoods to have lower crime rates, partly because they may employ effective defense strategies against crime.

Turning to sociological theories, it has been put forward that the prevalence of crime may be linked to a lack of upward mobility in society. Coser (1968, cited by Blau and Blau, 1982, p. 119) argues that people who perceive their poverty as permanent may be driven by hostile impulses rather than rational pursuit of their interests. Wilson and Daly (1997) hypothesize that sensitivity to inequality, especially by those at the bottom, leads to higher risk tactics, such as crime, when the expected payoffs from low-risk tactics are poor. If income inequality is correlated with social mobility, then these theories imply a higher prevalence of criminal behavior in more unequal areas.

Closely related to theories involving social mobility are those related to social disorganization and crime. In an influential paper, Merton (1938) proposes that the lack of upward mobility in a society, combined with a high premium on economic affluence results in anomie, a breakdown of standards and values. According to Merton, poverty or even “poverty in the midst of plenty” alone is not sufficient to induce high levels of crime. Only when their interaction with other interdependent social and cultural variables is considered, one can explain the association between crime and poverty.

The above theories, connecting crime to inequality, have spawned a large number of empirical studies. Most of these have estimated whether crime rates are affected by different measures of income inequality, using such measures as the Gini coefficient, the variance of log income, and different percentile quotients, like the 90/10-quotient. These different measures do however yield quite different conclusions in different studies. For example,

\(^3\) Hence, income inequality is supposed to be associated with crime levels via a relationship with the returns from crime and non-crime activities.
while certain studies using US data find a significant and positive relation between the Gini coefficient and crime rates (see Freeman, 1999, for an overview), Nilsson (2004) find no significant effects from the Gini coefficient on crime rates using Swedish data.

We believe that one explanation for the diverging results in the literature may be due to the use of an aggregate measure of income. It is a relatively old thought in economics, dating back at least to Friedman and Kuznets (1954), that an individual’s income in a given period can be divided into a permanent and a transitory component. Since changes in permanent and transitory income have different ramifications for the duration of inequality, it is quite likely that they will have different impacts on the crime rates. From the sociological theories related to above, it is clear that it is an individual’s permanent position in society that is the main factor affecting one’s decision to commit crime or not, not the individual’s transitory deviation from the permanent position. From the economic theories, it is however not clear whether it is inequality in permanent or in transitory income that matters. Using an aggregate measure of income, as the earlier studies have done, will however restrict inequality in permanent and transitory income to have the same impact.

To make the argument clear, consider the following model:

\[ y_{it} = u_i + \varepsilon_{it}, \quad (1) \]

where \( y_{it} \) is the log of total income in period \( t \) for individual \( i \), \( u_i \sim (\mu, \sigma_u^2) \) is permanent income for individual \( i \) which is assumed to be constant over the life-cycle and have a constant variance, and \( \varepsilon_{it} \sim iid(0, \sigma_\varepsilon^2) \) capture transitory stochastic deviations from permanent earnings. With this model, the cross-sectional variance of income in year \( t \) is:

\[ Var(y_{it}) = \sigma_u^2 + \sigma_\varepsilon^2, \quad (2) \]

that is, the variance of total income is the sum of the variance of permanent income and the variance of transitory income.

Equation (2) illustrates the potential pitfall of using total income. For example, suppose that changes in the variance of total income is used to study whether income inequality can explain differences in crime rates across different regions. If only permanent inequality affect crime rates, equation (2) shows that such a study has rather limited prospects of obtaining clear or systematic evidence. A region with both low inequality in permanent income and low
crime rates may have a large dispersion in transitory earnings, and hence a large cross-
sectional variance in income. The results may hence show that inequality has no, or even a
negative, effect on crime.

In the end, it is an empirical question whether inequality in permanent income has
another impact on crime rates than inequality in transitory income. Therefore, the aim with
this paper is to allow the two income components in equation (2) to have separate effects on
the crime rates.

III. Permanent and transitory income

An econometric model of income dynamics

Even though the permanent earnings model presented in the preceding section is intuitive,
previous studies have shown that additional features must be added in order to correctly
capture changes in permanent and transitory inequality; see the discussion in Baker and Solon
(2003) and the references therein. In particular, the permanent and transitory income
components should be allowed to vary with both time and age, and transitory shocks should
be allowed to last for several periods. The following model incorporates these features:

\[ y_{ita} = p_i u_{ita} + q_a e_{ita}, \]  

\[ u_{ita} = u_{ita-1} + r_a, \]  

\[ e_{ita} = \rho e_{ita-1} + \delta v_{ita-1} + \nu_a, \]

and where \( a \) denotes age. Beginning with the permanent component, \( p_i \) in equation (3) is a
year-specific factor loading that allows the permanent component to change over time.
Equation (4) models individual permanent income as a random walk where the innovation is
\( r_a \sim \text{iid}(0, \sigma^2_{r,a}) \). The innovation variance is allowed to take on one value up until age 34 and
another one thereafter. We also estimate the variance of an initial permanent shock that
capture the accumulation of individuals’ permanent shocks up to the age of 25 – the youngest
defined age in our sample – denoted $\sigma^2_{25}$. We found that not allowing for age variation in the
variance of permanent shocks produced noticeable different results as well as substantially
larger standard errors for the estimated parameters. By allowing for age variation we
recognize that younger individuals are more likely to be engaged in job-shopping and have no
job-securing tenure and are therefore more likely to experience larger permanent shocks to
income.$^4$

For the transitory component, $q_a$ in equation (3) is an age-specific parameter that
permits the magnitude of the transitory component to vary with age.$^5$ It is allowed to take on
three different values, one for individuals aged no more than 34 years, one for individuals
aged 35 to 44, and one for individuals aged 45 and above. Like for the permanent component
we found age variation to be important for the precision of the estimates. Equation (5) models
deviations from permanent income as a first order autoregressive moving average process
with a year-specific innovation $\nu_{it}$, which in turn has a year-specific variance, denoted $\sigma^2_{\nu}$.6

With the model in equations (3)–(5), the variance of log earnings for an individual no
more than 34 years old is

$$
Var(y_{it}) = p_i^2(\sigma^2_{25} + (a-25)\sigma^2_{r,26-34}) + q_a^2[\rho^2Var(\epsilon_{i,t-1}) + (2\rho\delta + \delta^2)\sigma^2_{\nu-1} + \sigma^2_{\nu}], \quad (6)
$$

where $\sigma^2_{r,26-34}$ is the variance of the random walk innovation for individuals aged 26–34. The
variance for an individual older than 34 years is

$$
Var(y_{it}) = p_i^2(\sigma^2_{25} + 9\sigma^2_{r,26-34} + (a-34)\sigma^2_{r,35-55}) + q_a^2[\rho^2Var(\epsilon_{i,t-1}) + (2\rho\delta + \delta^2)\sigma^2_{\nu-1} + \sigma^2_{\nu}], \quad (7)
$$

---

$^4$ A random walk specification is also used, among others, in Dickens (2000), Moffitt and Gottschalk (2002), and
Gustavsson (2004, 2007). We have also tried to fit the more general model of Baker and Solon (2003), which
incorporates both a random walk and a random growth model in the permanent component. For many of the
counties, this did however result in negative estimates of some of the variances or in large standard errors. One
potential explanation for this outcome is the smaller sample sizes that result from performing estimations at the
county level.

$^5$ In this model, the terms “permanent” and “transitory” are questionable. To be consistent with previous studies,
however, we use the term permanent for the non-mean reverting component and transitory for the mean
reverting component.

$^6$ The model in equations (3)–(5) is the most advanced model possible with our data. We have not applied
Newey’s (1985) specification test as the previous literature on income dynamics show that this test always (at
least in the studies that we are aware of) rejects the hypothesis that the right model is specified. The drawbacks
with this test for assessing the goodness of fit of models of income dynamics are further discussed in Baker
(1997) and Baker and Solon (2003). A general critique of tests such as that proposed in Newey (1985) is also
found in Leamer (1983).
where $\sigma^2_{r,35-55}$ is the variance of the random walk innovation for individuals aged 35–55.

Though equations (6) and (7) are complex, their implications for inequality are straightforward. An increase in the dispersion of either permanent or transitory income both generates increased cross-sectional income dispersion. The character of the change depends crucially, however, on which of these two components that changes. A persistent rise in the permanent component increases long-run inequality as the relative labor market advantage of workers with chronically high income is enhanced. An increase in the transitory component, without any change in the dispersion of permanent income, generates increased cross-sectional income dispersion by raising year-to-year income instability but with no change in long-term inequality. The reader is further referred to Baker and Solon (2003) for a detailed discussion about how models such as that in equations (3)–(5) should be interpreted and what they imply about individual earnings dynamics.

**Income data and estimation methodology**

To calculate permanent and transitory income inequality from 1974 to 2000 we use the register-based longitudinal database LINDA, constructed to be cross-sectionally representative of the Swedish population each year. The dataset is large; it contains 3.35 percent of the Swedish population each year corresponding to over 300,000 individuals. An attractive feature of the database is that attrition from the sample can be due only to death or migration. Information about individuals’ incomes comes from tax reports, so the income variable is free of the measurement errors that are common in survey data such as recall errors, rounding errors and top-coding.

A drawback with LINDA is that prior to 1983, income is only recorded for individuals with income higher than the threshold for being obligated to fill a tax report. The highest threshold is for 1974 with 21,213 real SEK (in year 2000 value). To get consistent samples over time, individuals with annual income below the 1974 threshold are excluded in any given year. This will exclude some of the unemployed individuals in each year, but only those with very long spells of joblessness. In fact, as the 1974 threshold approximately corresponds to a months pay for full time work, individuals who during two years have an uninterrupted unemployment spell for as long as 22 months can be retained in both these years depending on how the spell falls across the calendar year. That is, an individual who is unemployed from

---

7 The registers are maintained by Statistics Sweden; see Edin and Fredriksson (2000) for details.
8 This corresponds approximately to 1,500 GPB, 2,300 EUR, or 3000 USD
February in year one to November in year two can still be included in both these years as long as he works full time in January in year one and in December in year two.

The definition of an individual’s income used in the analysis is the log of total earnings from all jobs during a year (including sickness benefits); like Nilsson (2004) we use this income definition since it is the most consistent over time available in the LINDA database. Consequently, our inequality measures will to a large extent reflect inequality in labor market outcomes. However, labor earnings are the primary source of income for a majority of people and therefore most likely highly correlated with alternative measures of income.

To estimate the parameters of equations (3)-(5) we employ the minimum distance estimator described in Chamberlain (1984) and Abowd and Card (1989). This means that variances and auto-covariances of income constitute the dependent variable in the estimations. To calculate these we first, for each county and year 1974 to 2000, select all males 20 to 59 years old with income above the 1974 threshold; thus, the resulting panels are unbalanced as individuals may have missing values for some years. This design permits us to use the largest possible sample in the construction of each element of the auto-covariance matrices. The restriction to males is because the large changes in female labor force participation during the sample period would confound an analysis of female income (this restriction is standard in the literature on permanent and transitory inequality). The average year-specific sample size for each county is displayed in Figure 1. It is worth pointing out that the sample sizes for the counties in many cases actually match the sample size in the PSID used in the US studies by Haider (2000) and Moffitt and Gottschalk (1995, 2002), and in some cases are larger. Table 1 further presents, for each year, the total number of men in our pooled county samples, the total number of men aged 20–59 in the LINDA database and the percentage of these that are excluded due to earnings below the threshold.

---

9 See Edin and Fredriksson (2000) for a detailed description of how the earnings measure in LINDA is constructed.
10 Of course, other measures of income, preferable disposable family income, could yield different estimates in how the permanent and transitory variances have evolved over time. However, since such a measure is not available in LINDA, we focus on labor earnings solely since this measure is the most straightforward to interpret.
11 The county of Gotland is excluded due to the very small sample size and because most of the crimes are committed by tourists.
12 The results would be confounded because we estimate a model of income, not of entry and exit; appending a model of entry and exit is beyond the scope of this paper.
Figure 1: The average year-specific sample size for each county

<table>
<thead>
<tr>
<th>Year</th>
<th># men in the pooled county samples</th>
<th># men in LINDA aged 20–59</th>
<th>% excluded due to earnings&lt;21,213 SEK</th>
</tr>
</thead>
<tbody>
<tr>
<td>1974</td>
<td>67,537</td>
<td>72,517</td>
<td>6.87</td>
</tr>
<tr>
<td>1975</td>
<td>68,190</td>
<td>72,963</td>
<td>6.54</td>
</tr>
<tr>
<td>1976</td>
<td>68,813</td>
<td>73,347</td>
<td>6.18</td>
</tr>
<tr>
<td>1977</td>
<td>68,565</td>
<td>73,131</td>
<td>6.24</td>
</tr>
<tr>
<td>1978</td>
<td>68,226</td>
<td>73,222</td>
<td>6.82</td>
</tr>
<tr>
<td>1979</td>
<td>68,367</td>
<td>73,533</td>
<td>7.03</td>
</tr>
<tr>
<td>1980</td>
<td>68,064</td>
<td>73,374</td>
<td>7.24</td>
</tr>
<tr>
<td>1981</td>
<td>67,487</td>
<td>73,244</td>
<td>7.86</td>
</tr>
<tr>
<td>1982</td>
<td>67,058</td>
<td>73,114</td>
<td>8.28</td>
</tr>
<tr>
<td>1983</td>
<td>67,081</td>
<td>73,191</td>
<td>8.35</td>
</tr>
<tr>
<td>1984</td>
<td>67,007</td>
<td>73,569</td>
<td>8.92</td>
</tr>
<tr>
<td>1985</td>
<td>67,440</td>
<td>74,090</td>
<td>8.98</td>
</tr>
<tr>
<td>1986</td>
<td>67,774</td>
<td>74,809</td>
<td>9.40</td>
</tr>
<tr>
<td>1987</td>
<td>69,345</td>
<td>75,389</td>
<td>8.02</td>
</tr>
<tr>
<td>1988</td>
<td>70,222</td>
<td>76,179</td>
<td>7.82</td>
</tr>
<tr>
<td>1989</td>
<td>71,067</td>
<td>77,052</td>
<td>7.77</td>
</tr>
<tr>
<td>1990</td>
<td>71,551</td>
<td>78,205</td>
<td>8.51</td>
</tr>
<tr>
<td>1991</td>
<td>70,379</td>
<td>78,492</td>
<td>10.34</td>
</tr>
<tr>
<td>1992</td>
<td>68,375</td>
<td>79,035</td>
<td>13.49</td>
</tr>
<tr>
<td>1993</td>
<td>65,850</td>
<td>79,817</td>
<td>17.50</td>
</tr>
<tr>
<td>1994</td>
<td>65,763</td>
<td>80,776</td>
<td>18.59</td>
</tr>
<tr>
<td>1995</td>
<td>66,734</td>
<td>81,112</td>
<td>17.73</td>
</tr>
<tr>
<td>1996</td>
<td>66,376</td>
<td>81,272</td>
<td>18.33</td>
</tr>
<tr>
<td>1997</td>
<td>65,992</td>
<td>80,423</td>
<td>17.94</td>
</tr>
<tr>
<td>1998</td>
<td>67,378</td>
<td>80,896</td>
<td>16.71</td>
</tr>
<tr>
<td>1999</td>
<td>68,030</td>
<td>80,591</td>
<td>15.59</td>
</tr>
<tr>
<td>2000</td>
<td>69,434</td>
<td>80,840</td>
<td>14.11</td>
</tr>
</tbody>
</table>
Though it is not possible to clearly map out why an individual in LINDA has earnings below the tax report threshold – it could be because of unemployment, studies, early retirement, or some other reason – Table 2 indicates that the unemployment rate has some effect on the number of excluded individuals. Most notably, the increase in the number of excluded men in the early 1990s is likely a reflection of the substantial rise in the unemployment rate that occurred around this period; in short, unemployment in Sweden exhibited small fluctuations around 2 percent up to the late 1980s, markedly increased during the first years of the 1990s to be 8 percent in 1995, followed by a rebound to 4 percent in 2000. In the next subsection we further discuss how this may affect our inequality estimates.

In constructing the auto-covariance matrices for different cohorts we employ the methodology used by Moffitt and Gottschalk (1995). For each year, individuals are categorized into four 10-year age cohorts: 20-29, 30-39, 40-49, and 50-59. This level of disaggregation assures a minimum of 160 individuals in the calculations of the auto-covariances, though a majority of the calculations invoke a substantially larger sample. In year $t$ of the data, we divide the individuals into these four age groups and follow them through to year $t+1$, year $t+2$, etc until either the end of our data is reached (the year 2000) or until the age interval in question reaches beyond age 59. For example, the 40-49 cohort in 1974 can be followed through to 1984 when they are 50-59, but no further. For each of the four cohorts, the variance for the initial year, $t$, and the auto-covariances between year $t$ and each subsequent year are calculated. A fresh set of cohorts is begun in each year, so in year $t+1$ individuals are again divided into the four age groups 20-29, 30-39, etc and the variance for $t+1$ and the auto-covariances between year $t+1$ and each subsequent year are calculated. The end result of this procedure is 1004 unique elements in each county’s covariance matrix. This way of constructing the covariance matrix ensures that every individual’s element is included uniquely into one cell of the matrix, at the same time allowing us to maximize the number of elements since we can use a fully unbalanced panel with constant age intervals in each year.

Each county’s resulting auto-covariance matrix is used in the minimum distance estimation of equations (3)-(5). Basically, the implied variances and auto-covariances of the model are fitted to the corresponding empirical moments in the data by non-linear least squares. In these estimations the age of the cohort aged 20-29 is defined to be 25, the age of

---

13 The individuals in the age group 50-59 can not be followed at all. However their incomes are used to construct diagonal elements of the covariance matrix, i.e., to calculate the variances.
the cohort aged 30-39 is 35, and so forth. Appendix A contains a description of the estimation procedure.

Results

Here we will briefly describe the estimation results for three selected counties. Table B1 in Appendix B contains the parameter estimates and associated standard errors for the county of Stockholm (with the total number of included men \( n = 31,585 \)) together with results for the most northern and southern counties, Norrbotten (\( n = 4,841 \)) and Skåne (\( n = 17,724 \)) respectively.\(^{14}\) Based on the parameter estimates we can obtain age-specific estimates of permanent and transitory inequality.

Figure 1a: Permanent and transitory inequality in the county of Stockholm, 1974-2000

\(^{14}\) For corresponding estimates for the whole of Sweden during this period, see Gustavsson (2004, 2007).
Figure 1b: Permanent and transitory inequality in the county of Norrbotten, 1974-2000

Figure 1c: Permanent and transitory inequality in the county of Skåne, 1974-2000

Figure 1a-c contains the predicted evolution of permanent and transitory inequality from 1974 to 2000 for individuals defined to be 45 years old (ie in the age interval 40-49) together with the predicted variance from the full model (the variance of permanent income plus the variance of transitory income) and the actual variance for 45-year olds. The permanent and transitory variance both make up around half of the total variance in all three counties. One would perhaps expect the transitory variance to be smaller, but the use of an unbalanced panel design tends to boost the transitory variance; see Gustavsson (2004) for a comparison of results from unbalanced and balanced panel designs.\footnote{The higher transitory variance will not affect our final crime estimates as long as the changes over time are unaffected. Note also that there is no a priori reason to prefer a balanced panel since this design induces...} Unlike previous studies...
we also include immigrants, who are likely to have more transitory income due to higher unemployment rates.

Cross-sectional inequality among 45-year olds is largest in Stockholm and smallest in Norrbotten. The estimates of the permanent and transitory component are more irregular for Norrbotten, likely reflecting the smaller sample for this county (due to fewer people living there). There are year-to-year differences in the evolution of permanent and transitory variances across the counties up until the mid 1990s. The major movements are similar though, with only small changes up to the late 1980s and large increases during the deep Swedish recession in the first half of the 1990s. There are, however, larger differences from the mid 1990s as permanent inequality is constant in Stockholm but decreases in Skåne and Norrbotten. Transitory inequality also displays a stronger increase in Skåne and Norrbotten than in Stockholm during this latter period. Overall, there are some important cross-county differences in the evolution of inequality in permanent and transitory income – especially during the 1990s. This is important to note, since it constitutes the identifying variation for the income variables in the estimation of the crime equations.

Even though, as discussed in the previous subsection, the presence of the tax threshold excludes some of the individuals with longer spells of unemployment from our samples, Figures 1a–c strongly indicates that our estimates still capture earnings inequalities associated with unemployment. Most notably, both permanent and transitory inequalities increase in tandem with the 1990s rise in unemployment, and this also applies to most other counties. Since there were only minor changes in wage inequality during this period (Edin and Holmlund, 1995; Gustavsson, 2006), these changes must stem from increased dispersion in hours worked. The rise in transitory earnings inequality is then consistent with the fact that unemployment increased markedly for all groups on the Swedish labor market, see Holmlund (2006), whereas the increase in persistent inequality is consistent with the fact that groups with lower earnings were hit harder by the crisis, see Gustafsson and Palmer (2002).
One feature of our econometric model of income dynamics is that it results in age-specific estimates of permanent and transitory inequality – a consequence of the division of individuals into age cohorts, which in turn is needed to assure enough variances and covariances in the estimations. It is hence natural to ask which age we should use when we investigate the connection to crime rates. A first suggestion might be to use the estimates for the youngest age, i.e., for those in the age interval 20-29 (defined to be 25 years old in the estimations), since young individuals commit most crimes. However, our estimates show that the magnitudes and life-cycle patterns of permanent and transitory inequality differ across counties (different estimates of the random walk variances and of the $q_a$ parameters). To illustrate the meaning of this, Figure 2 graphs the life-cycle pattern of permanent inequality holding time constant, for the county of Stockholm, Norrbotten, and Skåne. The figure shows that inequality among 25-year olds actually is largest in Norrbotten, but also that this is a quickly passing state as the age-inequality pattern is much steeper in Stockholm and Skåne. Hence, even if inequality is low in Stockholm among individuals aged 25, those with low expected permanent income over the life-cycle will have much lower expected relative income in the future than the corresponding 25-year old individuals in Norrbotten. Since, according to economic theories along the lines of Becker (1968), expected future relative income is closely connected to the decision to commit crime, it is important to take account of the life-cycle differences across counties when estimating the connection between permanent inequality and crime rates.

---

16 Since, by construction, the proportional changes in inequality over time within a county are the same for all ages, we cannot include several different ages since these would be perfectly collinear.
To get measures of inequality that take account of the cross-county differences in life-cycle patterns, we compute the average permanent and transitory inequality over the life cycle for each county. That is, for a given year and county we calculate the permanent variance for each age (25, 26, and so forth up to 55), and then compute the average of these variances. The corresponding is done for transitory inequality. These averages of permanent and transitory inequality are the measures that we use in our estimated crime equations.

IV. Data and econometric specification of crime equation

When estimating the crime equations, we will use panel data from Swedish counties. In the panel, we observe the 20 counties over the time period 1974-2000. We will investigate the effects of income inequality on four different types of crime: total crime, shoplifting, auto theft, and burglary. Time-series graphs of these crime categories are available in Appendix B.

We assume that there is an underlying process that connects per capita crime rates, $C_{it}$, and the two measures of income inequality, $\text{PERM}_{it}$ and $\text{TRANS}_{it}$, in county $i = 1, \ldots, N$ at time $t = 1, \ldots, T$ such that

$$C_{it} = e^{\alpha \text{PERM}_{it} + \beta \text{TRANS}_{it}}$$

where $\text{PERM}_{it}$ is the estimated inequality in permanent income and $\text{TRANS}_{it}$ is the estimated inequality in transitory income and where we are interested in estimating the parameters $\alpha$ and $\beta$. Since crime rates are non-negative, the exponential form is suitable. Furthermore, for the exponential form, any changes are proportional to the crime rate, which seems more plausible than for example constant changes produced by a linear relation.

In order to avoid misspecifications due to omitted variables, we control for observable as well as unobservable variables that might explain the crime rate and that might be correlated with the two inequality measures. The observable variables, $x_{it}$, that we use are crime-specific clear-up rates (proxy for the probability of getting caught), unemployment rate, mean income measured in 100,000 SEK, share of men in the age interval 15-24, share of

---

17 The county of Gotland is excluded; see footnote 11.
18 The crime data is defined as the reported crime per 100,000 inhabitants and year, and it is collected from The Swedish National Council for Crime Prevention (BRÅ). With total crime we mean all reported crimes (not just the total of the crime categories that we use in this paper).
foreign citizens, and share of the population that is divorced. The unobservable variables are county-specific fixed effects, $f_{it}$, to control for unobserved variables that affects the crime rate and that stay constant over time for each county, time-specific fixed effects, $\lambda_t$, to control for unobserved macro-economic shocks that affect the crime rate in each county in the same way in a given year, and county-specific time trends, $\text{trend}_i$. The fixed county- and time-effects are allowed to be correlated with the observable variables. Furthermore, the crime rate can also be affected by disturbances, $\varepsilon_{it}$. Thus, we have the following relationship to be estimated for the crime rate

$$C_{it} = e^{\alpha \text{PERM}_{it} + \beta \text{TRANS}_{it} + \delta' x_{it} + f_{it} + \lambda_t + \text{trend}_i + \varepsilon_{it}}$$  \hspace{1cm} \text{(8)}$$

To estimate equation (8), we take the logarithm of it and use OLS to estimate the following familiar log-linear fixed effect model$^{20}$:

$$\ln(C_{it}) = \alpha \text{PERM}_{it} + \beta \text{TRANS}_{it} + \delta' x_{it} + f_{it} + \lambda_t + \text{trend}_i + \varepsilon_{it}$$  \hspace{1cm} \text{(9)}$$

Descriptive statistics for the variables used in the crime equations are provided in Table 2.

There are three things that we have to deal with when estimating equation (9). First, it is quite likely that the disturbances are heteroscedastic. Crimes are discrete events and the number of crimes committed is an integer. While this is not a problem for larger populations, for smaller populations it is, since the discrete nature of the crimes then will transfer to the crime rate, which is our dependent variable. For a population of 5,000, one additional crime corresponds to 20 crimes per 100,000 inhabitants. Since the precision of crime rate estimates as a consequence will depend on the population size, we cannot expect the variance of the regression errors to be homoscedastic, if we estimate equation (9) with common methods.

\footnote{These are control variables that are typically used when estimating crime equations; see eg, Raphael and Winter-Ebner (2001), Donohue and Levitt (2001) and Edmark (2005). We do not observe a crime-specific clear-up rate for shoplifting. In the empirical analysis, we use the total clear-up rate as a regressor in the equation for shoplifting.}

\footnote{The log-linear fixed effect model has been frequently used in papers estimating crime equations; see eg, Raphael and Winter-Ebner (2001) and the references cited therein. Also, some earlier studies have used weighted least squares (WLS). However, efficiency gains from WLS over OLS is only guaranteed if the conditional variance of the residuals is correctly specified, something that generally is very hard to do, and becomes even harder in our setting since we have generated regressors; see e.g. the discussion in Stock and Watson (2007). WLS also requires stronger assumptions for consistency; see Wooldridge (2002). Because of this, we focus on OLS with robust and bootstrapped standard errors. The interested reader is referred to the working paper version, Dahlberg and Gustavsson (2005), for results based on WLS.}
The smaller the population is, the larger is the variance.\textsuperscript{21} We take care of this problem by estimating robust standard errors. Second, if there is a serial correlation in the error process (that is, if the crime rates are serially correlated), the resulting standard errors are inconsistently estimated and may lead to severely biased estimates in small samples (see, eg Kezdi, 2002, and Bertrand et al., 2004). Therefore, we allow the errors to be correlated over time within each county.\textsuperscript{22} Third, since \textit{PERM}_i and \textit{TRANS}_i are estimated variables, their estimated standard errors might be biased. To correct for this, we present bootstrapped confidence intervals for these two variables.\textsuperscript{23}

\begin{table}
\centering
\caption{Descriptive statistics for the variables used in the crime equations} 
\begin{tabular}{lccc}
\hline
\textbf{Dependent variables} & \textbf{Mean} & \textbf{Std. Dev.} & \textbf{Min} & \textbf{Max} \\
\hline
Total crime & 9.34 & .21 & 8.91 & 9.98 \\
Shoplifting & 7.46 & .93 & 5.46 & 9.86 \\
Auto theft & 6.13 & .48 & 5.15 & 7.56 \\
Burglary & 7.23 & .26 & 6.64 & 7.88 \\
\hline
\textbf{Income variables} & & & & \\
Predict. variance & .350 & .06 & .21 & .53 \\
Gini & .27 & .03 & .21 & .35 \\
Variance & .37 & .06 & .23 & .55 \\
Permanent & .12 & .02 & .07 & .18 \\
Transitory & .23 & .05 & .13 & .35 \\
\hline
\textbf{Control variables} & & & & \\
Unemployed & .06 & .03 & .008 & .13 \\
Foreign & .04 & .02 & .01 & .10 \\
Divorced & .07 & .01 & .05 & .10 \\
Men 15-24 & .07 & .005 & .06 & .08 \\
Mean income & 1.76 & 0.13 & 1.42 & 2.33 \\
Clear-up, total & .21 & .04 & .12 & .32 \\
Clear-up, auto & .14 & .06 & .04 & .32 \\
Clear-up, burglary & .08 & .07 & .02 & .7 \\
\hline
\end{tabular}
\end{table}

\textit{Note:} We do not observe a crime-specific clear-up rate for shoplifting. In the empirical analysis, we use the total clear-up rate as a regressor in the equation for shoplifting. Mean income is measured in 100,000 SEK.

V. Results

For comparative reasons we start by using a traditional, aggregate, measure of income when calculating income inequality. Three measures are used: the predicted variance (which is the sum of the estimated variances in permanent and transitory income; cf equation (2)), the Gini

\textsuperscript{21} See Osgood (2000) for a discussion in a cross-sectional setting.
\textsuperscript{22} Technically, this is done in STATA by clustering on county.
\textsuperscript{23} We use 1000 bootstrap replications to estimate the confidence intervals.
coefficient, and the variance. The results of including these measures in equation (10) are presented in Table 3.

If it is the case that permanent income has an effect on crime while transitory income does not, then we might end up with the false result of insignificant effects when using an aggregate income measure. This is also what we get; none of the measures we use for income inequality enters significantly for any of the crime categories.24

Next we turn to the, for the purpose of this paper, more interesting question of whether the results in Table 3 change when we separate aggregate income into a permanent and a transitory part. The results are presented in Table 4. For each crime category, we present two different sets of estimates; in column (1) we present the results when we do not control for the crime-specific clear-up rate and in column (2) we present the results when we control for the clear-up rate. The reason for this division is the potential problems the endogeneity of the clear-up rate might cause in interpreting the coefficients for the permanent and transitory income variables.25 In parenthesis, we present the traditionally estimated standard errors. However, for the two inequality variables the estimated standard errors are biased since the variables are estimated. Therefore, we rely on bootstrap confidence interval when making inference for the inequality variables.26 95 and 99 percent bootstrap confidence intervals are presented within brackets.

From the results in Table 4, it is clear that it is important to decompose income. Starting with the inequality in permanent income, it turns out that this variable enters statistically significant in all crime equations. For total crime it is statistically significant at the one percent significance level (as a matter of fact, all 1000 bootstrap estimates are positive for this crime category), for shoplifting and auto theft it is significant at the five percent significance level, and for burglary the 95 percent bootstrap confidence interval never covers the zero when the clear-up rate is included in the model and it is significant at the ten percent

24 These results are in line with Nilsson (2004).
25 As can be seen from Table 4, the estimates for the two income variables are very similar in columns (1) and (2), indicating that the clear-up rate does not cause any problems in interpreting the coefficients for the income variables. We have also estimated all models with a one-year lag of the clear-up rates instead of contemporaneous clear-up rate (results not reported), which is an approach that sometimes is used to mitigate problems of endogeneity. This did not alter our conclusions for the income variables. Due to space constraints in Table 3, we did not report the results when the clear-up rates where included in those regressions. However, when the clear-up rates where included, we obtained very similar results for the income variables as those reported in Table 3.
26 In each bootstrap iteration, we first draw individuals with replacement from the income data base (ie, from the LINDA data base). Then we use the procedure presented in section 3 to estimate the county-specific inequalities in permanent and transitory income. Finally, we use the estimated county-specific inequalities in permanent and transitory income and estimate the crime equations. This procedure is repeated 1000 times, providing us with 1000 estimates on the coefficients for the two income variables.
significance level when the clear-up rate is not included (the latter confidence interval is not
reported in the table). Turning to the inequality in transitory income, this variable never
enters significantly; the zero is covered by all confidence intervals at conventional
significance levels (at the ten percent significance level or higher). It hence seems like it is
inequality in permanent income – and not in transitory income - that is important in
determining crime, at least the crime categories considered here.

The results for the other variables in Table 1 and 2 are in line with those in previous
studies - with the noticeable exception of the unemployment rate which enters insignificantly
in all specifications but one. This result for the unemployment rate is perhaps not surprising
as compared to the bulk of the earlier studies aiming at estimating the effect of unemployment
on crime (for a discussion on this issue, see Raphael and Winter-Ebmer, 2001). However, it is
rather surprising given the results in more recent work, such as Raphael and Winter-Ebmer
(2001), Edmark (2005), and Öster and Agell (2007). It is especially surprising given that we
are using the same type of data as in Edmark (county-level data for Sweden). There might
however be several explanations for the diverging results. One explanation for the different
results could be the disaggregation of the income variable into its two components, as done in
this paper. However, from the results in Table 1 where total income is used, this does not
seem to be the explanation. Compared to us, the baseline specification in Edmark (i) uses a
shorter time period (1988-1999), (ii) uses a log-log specification, (iii) have no county-specific
time trends, and (iv) do not cluster the standard errors on cross-sectional unit. Furthermore,
she examines more crime categories than we do. The crimes we have in common are
shoplifting, auto theft and burglary. Edmark finds that the coefficient for the unemployment
variable is significant for burglary and auto theft. In order to try to reconcile our results with
those of Edmark, we conducted some further analysis for the time period used by Edmark.
From the results (not reported) it seems like the results for the unemployment coefficient are
sensitive to the inclusion of county-specific time trends and to clustering on cross-sectional
unit. The latter result is in some sense in line with the sensitivity results presented by Edmark,

---

27 For the total crime category, the coefficients for permanent inequality are not covered by any of the bootstrap
confidence intervals, something that could indicate a bias in the estimates for this crime category. One reason for
this could be that a richer model is needed to explain the evolution of all types of crimes (in which, for example,
both all property crimes and all violent crimes are included). It is however reassuring that the mean bootstrap
value is lower than our main estimate, since this indicates that we have a downward bias in the estimated impact
of permanent inequality on total crime rates; see e.g. Cameron and Trivedi (2005) for a discussion of these
issues.

28 As a sensitivity analysis we have also conducted estimations based on an age-weighted average of permanent
and transitory inequality (ie, when averaging the inequalities over the different ages, we calculate a weighted
average where the share of individuals at each age constitute the weights; in the baseline analysis, we used
straight averages). This does not alter the conclusions.
where she gets less significant results when including county-specific time trends (c.f. her Table 4).

VI. Conclusions

Earlier studies on income inequality and crime have used inequality in total income or total earnings as explanatory variable. However, from sociological theories on inequality and crime, it is rather an individual’s permanent position in society that is the main factor affecting one’s decision to commit crime or not, not the individual’s transitory deviation from the permanent position, implying that it is the inequality in permanent income, and not in transitory income, that is the important determinant for crime. Although less clear, this can also be the prediction from economic theories. Using an aggregate measure of income, as the earlier studies have done, will however restrict inequality in permanent and transitory income to have the same impact, making it difficult to obtain systematic evidence if only one of the components matter (or if the two components matter to different degrees).

We have in this paper investigated whether the effects from inequality in permanent income on crime differ from the effects from inequality in transitory income on crime. To that end, we started out by estimating, using very rich income data from Sweden, region-specific inequality in permanent and transitory income. Then we used the two estimated income components as explanatory variables in four different crime equations; total crime, shoplifting, auto theft, and burglary.

The results indicate that it is crucially important to separate the two effects; while an increase in the inequality in permanent income yields a positive and significant effect on total crimes and the three different property crimes, an increase in the inequality in transitory income has no significant effect on any type of crime. Using a traditional, aggregate, measure of income yields insignificant effects on crime.
### TABLE 3

*Estimates when using inequality in aggregate measure of income*

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predict. variance</td>
<td>0.213</td>
<td></td>
<td></td>
<td>0.739</td>
<td></td>
<td></td>
<td>0.642</td>
<td></td>
<td></td>
<td>0.114</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.245)</td>
<td></td>
<td></td>
<td>(0.681)</td>
<td></td>
<td></td>
<td>(0.729)</td>
<td></td>
<td></td>
<td>(0.531)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gini</td>
<td>0.355</td>
<td>4.597</td>
<td></td>
<td>-0.433</td>
<td>-0.792</td>
<td></td>
<td>-0.433</td>
<td></td>
<td></td>
<td>-0.792</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.817)</td>
<td>(3.174)</td>
<td></td>
<td>(0.817)</td>
<td>(3.174)</td>
<td></td>
<td>(3.086)</td>
<td></td>
<td></td>
<td>(1.304)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td>0.267</td>
<td></td>
<td></td>
<td>0.984</td>
<td></td>
<td></td>
<td>0.204</td>
<td></td>
<td></td>
<td>0.129</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.210)</td>
<td></td>
<td></td>
<td>(0.814)</td>
<td></td>
<td></td>
<td>(0.809)</td>
<td></td>
<td></td>
<td>(0.576)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unemployed</td>
<td>1.562</td>
<td>1.564</td>
<td>-1.131</td>
<td>-1.420</td>
<td>-1.141</td>
<td>2.815</td>
<td>2.951</td>
<td>2.890</td>
<td>0.547</td>
<td>0.635</td>
<td>0.550</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.978)</td>
<td>(0.967)</td>
<td>(2.134)</td>
<td>(2.150)</td>
<td>(2.136)</td>
<td>(2.249)</td>
<td>(2.156)</td>
<td>(1.227)</td>
<td>(1.262)</td>
<td>(1.224)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean income</td>
<td>0.292</td>
<td>0.244</td>
<td>0.337</td>
<td>-0.679</td>
<td>-0.270</td>
<td>-0.489</td>
<td>-0.217</td>
<td>-0.623</td>
<td>-0.429</td>
<td>0.547</td>
<td>0.350</td>
<td>0.563</td>
</tr>
<tr>
<td></td>
<td>(0.244)</td>
<td>(0.268)</td>
<td>(0.237)</td>
<td>(0.740)</td>
<td>(0.763)</td>
<td>(0.698)</td>
<td>(0.688)</td>
<td>(0.628)</td>
<td>(0.728)</td>
<td>(0.632)</td>
<td>(0.577)</td>
<td>(0.672)</td>
</tr>
<tr>
<td>Observations</td>
<td>540</td>
<td>540</td>
<td>540</td>
<td>540</td>
<td>520</td>
<td>520</td>
<td>540</td>
<td>540</td>
<td>540</td>
<td>540</td>
<td>540</td>
<td>540</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.95</td>
<td>0.94</td>
<td>0.95</td>
<td>0.97</td>
<td>0.97</td>
<td>0.97</td>
<td>0.92</td>
<td>0.92</td>
<td>0.92</td>
<td>0.91</td>
<td>0.91</td>
<td>0.91</td>
</tr>
</tbody>
</table>

*Notes: Robust standard errors are presented in parentheses.*

Time dummies, county-specific fixed effects and county-specific time trends are included in all specifications. Clustering is made on counties (allowing for autocorrelation in the residuals).

* significant at 10%; ** significant at 5%; *** significant at 1%
### TABLE 4

Estimates when using inequality in permanent and transitory income

<table>
<thead>
<tr>
<th></th>
<th>Total crime</th>
<th>Shoplifting</th>
<th>Auto theft</th>
<th>Burglary</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Permanent</td>
<td>1.481</td>
<td>1.661</td>
<td>2.353</td>
<td>2.363</td>
</tr>
<tr>
<td>[95% bootstrap CI]</td>
<td>[0.43 1.14]</td>
<td>[0.49 1.23]</td>
<td>[0.24 2.11]</td>
<td>[0.26 2.09]</td>
</tr>
<tr>
<td>[99% bootstrap CI]</td>
<td>[0.10 1.37]</td>
<td>[0.12 1.44]</td>
<td>[-0.55 2.77]</td>
<td>[-0.54 2.74]</td>
</tr>
<tr>
<td>Transitory</td>
<td>-0.087</td>
<td>-0.116</td>
<td>0.425</td>
<td>0.288</td>
</tr>
<tr>
<td>[95% bootstrap CI]</td>
<td>[-0.31 0.19]</td>
<td>[-0.34 0.18]</td>
<td>[-0.49 0.85]</td>
<td>[-0.56 0.78]</td>
</tr>
<tr>
<td>[99% bootstrap CI]</td>
<td>[-0.53 0.31]</td>
<td>[-0.56 0.31]</td>
<td>[-0.96 1.18]</td>
<td>[-1.07 1.12]</td>
</tr>
<tr>
<td>Clear-up rate</td>
<td>0.054</td>
<td>2.395</td>
<td>-0.619</td>
<td>-0.004</td>
</tr>
<tr>
<td></td>
<td>(0.249)</td>
<td>(0.624)**</td>
<td>(0.285)**</td>
<td>(0.081)</td>
</tr>
<tr>
<td>Unemployed</td>
<td>1.300</td>
<td>1.399</td>
<td>-1.428</td>
<td>-0.020</td>
</tr>
<tr>
<td></td>
<td>(0.871)</td>
<td>(0.753)*</td>
<td>(2.110)</td>
<td>(2.118)</td>
</tr>
<tr>
<td>Foreign</td>
<td>6.435</td>
<td>5.831</td>
<td>3.419</td>
<td>3.258</td>
</tr>
<tr>
<td>Mean income</td>
<td>0.338</td>
<td>0.347</td>
<td>-0.595</td>
<td>-0.273</td>
</tr>
<tr>
<td></td>
<td>(0.246)</td>
<td>(0.253)</td>
<td>(0.754)</td>
<td>(0.723)</td>
</tr>
<tr>
<td>Observations</td>
<td>540</td>
<td>520</td>
<td>520</td>
<td>520</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.95</td>
<td>0.95</td>
<td>0.97</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Notes: Robust standard errors are presented in parentheses. Time dummies, county-specific fixed effects and county-specific time trends are included in all specifications. Clustering is made on counties (allowing for autocorrelation in the residuals).

* significant at 10%; ** significant at 5%; *** significant at 1%
References


Appendix A: The minimum distance estimation

The parameters of the model in equations (6)-(8) are estimated by applying the minimum distance estimator of Chamberlain (1984). Specifically, let $C_b$ contain the distinct elements of the population auto-covariance matrix of $y_{itb}$ for cohort $b$ and let $C$ be an aggregate vector stacked with the $C_b$ vectors. Let the vector $\theta$ contain all the parameters of our model and let $C = f(\theta)$ express the model’s moment restrictions. Our model then implies that the general variance element in $C$ is

$$Var(y_{it}) = p_t^2(\sigma_u^2 + \sum_a \sigma_{ra}^2) + q_a^2[\rho^2Var(\varepsilon_{i,t-1}) + (2\rho\delta + \delta^2)\sigma_{\eta_{i-1}}^2 + \sigma_{\eta_i}^2],$$  \hspace{1cm} (A1)

and that the general auto-covariance element for years $t$ and $t-s$ is

$$Cov(y_{it}, y_{i,a,t-s,t-s}) = p_t p_{t-s} (\sigma_u^2 + \sum_{a,s} \sigma_{rs}^2) + q_a q_{a-s}[\rho^2Var(\varepsilon_{i,t-s}) + \delta \rho^{s-1} \sigma_{\eta_{i-s}}^2].$$ \hspace{1cm} (A2)

The vector $C$ is estimated by the sample counterpart $\hat{C}$, and $\hat{\theta}$ is chosen to minimize a distance function

$$D = (\hat{C} - f(\hat{\theta}))'W(\hat{C} - f(\hat{\theta})), $$ \hspace{1cm} (A3)

where $W$ is a positive definite weighting matrix.

The asymptotically optimal choice of $W$ is the inverse of a matrix that consistently estimates the covariance matrix of $C$. However, Altonji and Segall (1996) and Clark (1996) provide Monte Carlo evidence of potentially serious finite sample bias in the estimate of $\theta$ using this approach. We therefore follow the practice of the most recent literature and use the identity matrix as the weighting matrix. This “equally weighted minimum distance estimation” amounts to using non-linear least squares to fit $f(\hat{\theta})$ to $\hat{C}$.

As outlined in Chamberlain (1984), auto-correlation and heteroskedasticity robust standard errors for $\hat{\theta}$ are obtained from the formula

$$G'G)^{-1}G'VGG'(G')^{-1},$$ \hspace{1cm} (A4)

---

29 This bias arises because of correlated sampling errors in the second and fourth moments of income.
where $\mathbf{G}$ is the gradient matrix $\partial f(\mathbf{0})/\partial \mathbf{\theta}$ evaluated at $\hat{\mathbf{\theta}}$ and $\mathbf{V}$ is a block diagonal matrix where the diagonal contains the estimated covariance matrices of each $\hat{\mathbf{C}}$, vectors.

Appendix B

Table B1

Estimates of income dynamics for three Swedish counties

<table>
<thead>
<tr>
<th>Permanent component</th>
<th>Stockholm</th>
<th>Norrbotten</th>
<th>Skåne</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma^2_\alpha$</td>
<td>0.029 (0.003)</td>
<td>0.034 (0.007)</td>
<td>0.028 (0.003)</td>
</tr>
<tr>
<td>$\sigma^2_{2.26-34}$</td>
<td>0.011 (0.001)</td>
<td>0.003 (0.001)</td>
<td>0.006 (0.001)</td>
</tr>
<tr>
<td>$\sigma^2_{2.35-55}$</td>
<td>0.002 (0.0002)</td>
<td>0.002 (0.0004)</td>
<td>0.003 (0.0004)</td>
</tr>
<tr>
<td>$p_{14}$</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>$p_{15}$</td>
<td>0.947 (0.045)</td>
<td>0.929 (0.123)</td>
<td>1.005 (0.064)</td>
</tr>
<tr>
<td>$p_{16}$</td>
<td>0.964 (0.045)</td>
<td>1.109 (0.137)</td>
<td>1.053 (0.065)</td>
</tr>
<tr>
<td>$p_{17}$</td>
<td>0.958 (0.043)</td>
<td>1.063 (0.131)</td>
<td>1.070 (0.065)</td>
</tr>
<tr>
<td>$p_{18}$</td>
<td>0.948 (0.042)</td>
<td>1.026 (0.122)</td>
<td>1.026 (0.062)</td>
</tr>
<tr>
<td>$p_{19}$</td>
<td>0.958 (0.041)</td>
<td>1.071 (0.123)</td>
<td>1.054 (0.062)</td>
</tr>
<tr>
<td>$p_{20}$</td>
<td>0.953 (0.041)</td>
<td>1.105 (0.125)</td>
<td>1.056 (0.062)</td>
</tr>
<tr>
<td>$p_{21}$</td>
<td>0.941 (0.039)</td>
<td>1.060 (0.118)</td>
<td>1.037 (0.060)</td>
</tr>
<tr>
<td>$p_{22}$</td>
<td>0.967 (0.040)</td>
<td>1.185 (0.130)</td>
<td>1.059 (0.061)</td>
</tr>
<tr>
<td>$p_{23}$</td>
<td>0.938 (0.039)</td>
<td>1.154 (0.125)</td>
<td>1.046 (0.060)</td>
</tr>
<tr>
<td>$p_{24}$</td>
<td>0.948 (0.039)</td>
<td>1.154 (0.122)</td>
<td>1.043 (0.059)</td>
</tr>
<tr>
<td>$p_{25}$</td>
<td>0.945 (0.038)</td>
<td>1.131 (0.119)</td>
<td>1.058 (0.059)</td>
</tr>
<tr>
<td>$p_{26}$</td>
<td>0.960 (0.039)</td>
<td>1.045 (0.112)</td>
<td>1.057 (0.059)</td>
</tr>
<tr>
<td>$p_{27}$</td>
<td>0.993 (0.040)</td>
<td>1.096 (0.116)</td>
<td>1.005 (0.056)</td>
</tr>
<tr>
<td>$p_{28}$</td>
<td>0.963 (0.038)</td>
<td>1.016 (0.107)</td>
<td>1.014 (0.056)</td>
</tr>
<tr>
<td>$p_{29}$</td>
<td>0.937 (0.038)</td>
<td>0.961 (0.102)</td>
<td>1.023 (0.057)</td>
</tr>
<tr>
<td>$p_{30}$</td>
<td>1.004 (0.040)</td>
<td>1.069 (0.112)</td>
<td>1.075 (0.059)</td>
</tr>
<tr>
<td>$p_{31}$</td>
<td>1.088 (0.043)</td>
<td>1.126 (0.117)</td>
<td>1.109 (0.061)</td>
</tr>
<tr>
<td>$p_{32}$</td>
<td>1.129 (0.044)</td>
<td>1.229 (0.126)</td>
<td>1.156 (0.063)</td>
</tr>
<tr>
<td>$p_{33}$</td>
<td>1.199 (0.046)</td>
<td>1.161 (0.119)</td>
<td>1.156 (0.063)</td>
</tr>
<tr>
<td>$p_{34}$</td>
<td>1.215 (0.047)</td>
<td>1.298 (0.130)</td>
<td>1.193 (0.064)</td>
</tr>
<tr>
<td>$p_{35}$</td>
<td>1.169 (0.044)</td>
<td>1.315 (0.130)</td>
<td>1.188 (0.063)</td>
</tr>
<tr>
<td>$p_{36}$</td>
<td>1.171 (0.043)</td>
<td>1.326 (0.131)</td>
<td>1.193 (0.063)</td>
</tr>
<tr>
<td>$p_{37}$</td>
<td>1.179 (0.043)</td>
<td>1.330 (0.130)</td>
<td>1.152 (0.061)</td>
</tr>
<tr>
<td>$p_{38}$</td>
<td>1.171 (0.043)</td>
<td>1.266 (0.124)</td>
<td>1.134 (0.060)</td>
</tr>
<tr>
<td>$p_{39}$</td>
<td>1.196 (0.043)</td>
<td>1.099 (0.107)</td>
<td>1.149 (0.060)</td>
</tr>
<tr>
<td>$p_{40}$</td>
<td>1.200 (0.043)</td>
<td>1.222 (0.118)</td>
<td>1.128 (0.059)</td>
</tr>
</tbody>
</table>

Transitory component

<p>| $q_{25-34}$ | 1.000 | 1.000 | 1.000 |
| $q_{35-44}$ | 0.723 (0.006) | 0.689 (0.014) | 0.690 (0.007) |
| $q_{45-55}$ | 0.684 (0.006) | 0.621 (0.013) | 0.626 (0.007) |</p>
<table>
<thead>
<tr>
<th>Year</th>
<th>$\rho$</th>
<th>$\delta$</th>
<th>$\sigma^2_\gamma$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1974</td>
<td>0.345 (0.012)</td>
<td>0.297 (0.025)</td>
<td>0.311 (0.016)</td>
</tr>
<tr>
<td>1975</td>
<td>0.286 (0.012)</td>
<td>0.298 (0.028)</td>
<td>0.219 (0.013)</td>
</tr>
<tr>
<td>1976</td>
<td>0.235 (0.011)</td>
<td>0.189 (0.025)</td>
<td>0.206 (0.014)</td>
</tr>
<tr>
<td>1977</td>
<td>0.207 (0.011)</td>
<td>0.222 (0.026)</td>
<td>0.175 (0.013)</td>
</tr>
<tr>
<td>1978</td>
<td>0.189 (0.010)</td>
<td>0.185 (0.023)</td>
<td>0.190 (0.013)</td>
</tr>
<tr>
<td>1979</td>
<td>0.169 (0.010)</td>
<td>0.157 (0.021)</td>
<td>0.182 (0.013)</td>
</tr>
<tr>
<td>1980</td>
<td>0.183 (0.010)</td>
<td>0.164 (0.021)</td>
<td>0.209 (0.013)</td>
</tr>
<tr>
<td>1981</td>
<td>0.186 (0.010)</td>
<td>0.199 (0.022)</td>
<td>0.213 (0.014)</td>
</tr>
<tr>
<td>1982</td>
<td>0.186 (0.010)</td>
<td>0.180 (0.023)</td>
<td>0.208 (0.013)</td>
</tr>
<tr>
<td>1983</td>
<td>0.217 (0.010)</td>
<td>0.221 (0.024)</td>
<td>0.212 (0.014)</td>
</tr>
<tr>
<td>1984</td>
<td>0.230 (0.011)</td>
<td>0.221 (0.024)</td>
<td>0.238 (0.014)</td>
</tr>
<tr>
<td>1985</td>
<td>0.212 (0.011)</td>
<td>0.230 (0.024)</td>
<td>0.218 (0.014)</td>
</tr>
<tr>
<td>1986</td>
<td>0.200 (0.011)</td>
<td>0.256 (0.027)</td>
<td>0.218 (0.015)</td>
</tr>
<tr>
<td>1987</td>
<td>0.204 (0.011)</td>
<td>0.221 (0.023)</td>
<td>0.226 (0.014)</td>
</tr>
<tr>
<td>1988</td>
<td>0.235 (0.011)</td>
<td>0.217 (0.025)</td>
<td>0.234 (0.014)</td>
</tr>
<tr>
<td>1989</td>
<td>0.235 (0.011)</td>
<td>0.247 (0.027)</td>
<td>0.233 (0.014)</td>
</tr>
<tr>
<td>1990</td>
<td>0.259 (0.012)</td>
<td>0.249 (0.028)</td>
<td>0.229 (0.015)</td>
</tr>
<tr>
<td>1991</td>
<td>0.274 (0.013)</td>
<td>0.274 (0.030)</td>
<td>0.276 (0.016)</td>
</tr>
<tr>
<td>1992</td>
<td>0.287 (0.014)</td>
<td>0.306 (0.033)</td>
<td>0.296 (0.017)</td>
</tr>
<tr>
<td>1993</td>
<td>0.309 (0.015)</td>
<td>0.320 (0.033)</td>
<td>0.329 (0.018)</td>
</tr>
<tr>
<td>1994</td>
<td>0.325 (0.016)</td>
<td>0.303 (0.035)</td>
<td>0.311 (0.019)</td>
</tr>
<tr>
<td>1995</td>
<td>0.321 (0.015)</td>
<td>0.261 (0.033)</td>
<td>0.317 (0.018)</td>
</tr>
<tr>
<td>1996</td>
<td>0.321 (0.015)</td>
<td>0.316 (0.036)</td>
<td>0.316 (0.019)</td>
</tr>
<tr>
<td>1997</td>
<td>0.340 (0.015)</td>
<td>0.305 (0.037)</td>
<td>0.377 (0.020)</td>
</tr>
<tr>
<td>1998</td>
<td>0.353 (0.015)</td>
<td>0.379 (0.041)</td>
<td>0.351 (0.020)</td>
</tr>
<tr>
<td>1999</td>
<td>0.357 (0.015)</td>
<td>0.369 (0.039)</td>
<td>0.376 (0.019)</td>
</tr>
<tr>
<td>2000</td>
<td>0.346 (0.014)</td>
<td>0.352 (0.035)</td>
<td>0.360 (0.018)</td>
</tr>
</tbody>
</table>

Note: The estimated model is outlined in equations (6)-(8). The estimates for each county are based on 1004 variances and auto-covariances. Heteroskedasticity and auto-correlation robust standard errors are in parentheses.
Appendix C

Time series graphs of total crime, shoplifting, auto theft, and burglary, for the period 1974-2000 (average reported crime rates per 100,000 inhabitants).

Figure C1: Average reported crime rates (per 100’ inhabitants) for the four crime categories, 1974-2000