Testing for Correlated Heterogeneity: Theory and Evidence on Couples' Incomes

Stephen H. Shore

January 2, 2008

Abstract

Identifying latent heterogeneity (heterogeneity in unobserved parameters that affect the distribution of an observed random variable) is a longstanding problem in economics. For example, some people may face income changes with larger variances than others, though we observe only realized income changes and not the variance parameters and shocks that generate them. Unfortunately, it is not possible to identify parameter heterogeneity affecting a single random variable (e.g., an individual-specific variance of income changes) without knowing the distribution of shocks, even in a well-behaved linear setting.

This paper shows that more headway can be made for correlated heterogeneity, when unobserved and potentially heterogeneous parameters affect the joint distribution of two variables. Dependence between two uncorrelated variables (because their second moments are correlated) can be interpreted as correlated parameter heterogeneity, regardless of the distribution of shocks. This paper introduces a "wife-swap bootstrap" test that extends this test – among others – to panel data. Correlated heterogeneity could result from heterogeneity in covariance parameters or from correlated heterogeneity in variance parameters. As with parameter heterogeneity affecting a single variable, it is possible to obtain lower bounds on each type of correlated heterogeneity affecting a pair of variables with additional covariates or panel data. Unlike the single-variable case, a lower bound on correlated variances implies an upper bound on covariance heterogeneity. The joint distribution of couples' income changes are used as an illustration.

Contents

List of Figures

List of Tables

1 Introduction

Identifying latent heterogeneity (heterogeneity in unobserved parameters that affect the distribution of an observed variable) is a longstanding problem in economics. For example, some people may face income changes with larger variances than others (Meghir and Pistaferri, 2004; Alvarez, Browning, and Ejrnaes, 2001); an econometrician cannot observe these variance parameters directly, but must infer them with considerable error from realized income shocks. Identifying this latent heterogeneity is important for two reasons. First, once an econometrician identifies how much variation in an observed variable reflects parameter heterogeneity, she can correct for the attenuation bias caused when that variable proxies for the parameter as a right-hand-side variable in a regression. For example, comparing the consumption or saving behavior of individuals with high and low variaces of income shocks provides a test of precautionary saving and buffer stock models (Carroll and Samwick, 1997), but only if we can extract true variation across individuals in the variance of their income shocks from estimation error. Second, welfare calculations with heterogeneity may differ from those under a representative agent framework. For example, differences in the variance of income shocks may be due to selection on the basis of risk-aversion (Browning and Lusardi, 1996). In this case, the welfare loss from income shocks will be substantially lower than under representative-agent models with average risk aversion and an average variance of income shocks (Schulhofer-Wohl, 2006). Without a measure of the heterogeneity in income risk across individuals (as in Jensen and Shore (2007)), such a calculation is impossible.

In general, it is not possible to use an observed variable to test for heterogeneity in an unobserved parameter, unless one is willing to assume the distribution of shocks to the variable conditional on the parameter. When observation-specific parameter estimates vary across observations, is this variation due to signal or noise? To heterogeneity across observations in parameter values or the large magnitude of realized shocks conditional on those values? When the parameter in question is a variance, is a fat-tailed distribution of realized values the result of heterogeneity in variance parmeters (signal) or fat-tailed shocks conditional on a constant variance (noise)? There is no way to resolve these questions without additional information.

This paper shows that more headway can be made in the two-variable case. Consider two variables that are unconditionally uncorrelated (their first moments are uncorrelated) but not independent (because their second moments are not). Assume a straightforward linear structure, so that common and idiosyncratic shocks are i.i.d., affine functions of unobserved parameters. In this case, correlated second moments can be decomposed into heterogeneity across observations in covariance parameters (heterogeneity in the coefficient on the common shock, signal) or correlated differences across observations in the two variables' variance parameters conditional on that covariance (correlated coefficients on idiosyntratic shocks, correlated noise). These are both types of heterogeneity, so that rejecting independence by finding correlated second moments can be interpreted as finding at least one of these types of heterogeneity. I jointly refer to these as "correlated heterogeneity." To identify correlated heterogeneity with this test, the two variables must be unconditionally uncorrelated.

This is a strong but very testable assumption. In the analogous single-variable case, both fat-tails and variance heterogeneity could explain large fourth moments. By contrast, when two variables are unconditionally uncorrelated, fat-tails have no impact on the covariance of their second moments, analogous to the fourth moment in the single-variable case.

As in the single-variable heterogeneity case, in the two-variable case additional information is needed to separate covariance heterogeneity from correlated variances. Panel data can be used to identify a lower bound on covariance heterogeneity and on correlated variances. Alternatively, additional covariates can also be used to identify a lower bound on covariance heterogeneity. Each of these techniques has a single-variable variance analog, placing a lower bound on heterogeneity in variance. However, identifying lower bounds on covariance heterogeneity (coupled with an approximate guess about tail fatness) places an upper bound on correlated variances, and *vise versa*. By contrast, in the single-variable variance heterogeneity case, a small mis-specification in tail fatness can lead to dramatic mis-measurement in variance heterogeneity.

These ideas are applied to the case of couples' joint income processes using data from the Panel Study of Income Dynamics. In the data, wives' income changes are approximately uncorrelated with their husbands' income changes. However, they are not independent, as couples' squared income changes are unconditionally correlated. This is equivalent to saying that the product of couples' income changes has a higher sample variance than would be expected if incomes evolved independently, or that large changes in husbands' and wives' incomes tend to coincide. I use panel data to place a lower bound on correlated volatility; wives whose income shocks have large variances tend to be married to husbands whose income shocks also have large variances. Panel data also provides a lower bound on covariance heterogeneity, so that some couples' incomes move together while other couples' incomes move in opposite directions. Putting these together (and assuming correlated shocks are roughly normal) places an upper bound on covariance heterogeneity. Of the variation in the product of couples' contemporaneous income changes, 66 percent of it can be attibuted to realized shocks, what would be expected even if couples' incomes moved independently. The heterogeneity in covariance parameters is only between 2 percent (the lower bound on the reliability ratio) and 11 percent (the upper bound) of this variance. These bounds covariance heterogeneity are used in Shore (2007) to correct for the attenuation bias caused when a noisy estimate of covariance is used as a right-hand-side variables.

1.1 Setup

Consider two variables, x_i and y_i, that are mutually independent across observations, i. In the example used in this paper, x_i is the change in "excess" log income for a wife in couple i and y_i is the change in "excess" log income for her husband. The choice of who is x and who is y is meant to be easy to remember; x and y represent the second sex chromosome (XX for women and XY for men). The word "excess" (described in detail in Section 4.1) implies that any aggregate or predictable changes to income have been removed, so that x_i and y_i are residuals and therefore unconditionally mean zero by construction. The key assumptions needed for the results to follow are:

1. **x_i and y_i can be written as the linear combination of two shocks**, where one of these (e_{ci}) is shared in common and the other (e_{xi} and e_{yi} for x_i and y_i, respectively) is not. This can be written as

$$\begin{aligned} x_i &= \sigma_{xi}\left(\sqrt{1-r_{xi}^2}e_{xi} + r_{xi}e_{ci}\right); \qquad (1) \\ y_i &= \sigma_{yi}\left(\sqrt{1-r_{yi}^2}e_{yi} + r_{yi}e_{ci}\right). \end{aligned}$$

   e_{xi}, e_{yi}, and e_{ci} are assumed to be independent of one another (and across i, though this assumptions can be relaxed), with variances normalized to one for all i.[1] $\theta_i \equiv (\sigma_{xi}, \sigma_{yi}, r_{xi}, r_{yi})$, the variance-covariance parameters for an observation are viewed as stochastic. Unless otherwise noted, references to conditional expectations mean expectation conditional on θ_i.

2. **e_{xi}, e_{yi}, and e_{ci} are independent of θ_i**, and independent of any additional covariates or calendar time variables that may be introduced later. For simplicity of exposition and equivalently for the econometrician, the shape of the distribution of e_{xi} is assumed to be identical for all i, as for e_{yi}, and e_{ci}. However, the distributions of e_{xi}, e_{yi}, and e_{ci} need not be identical to one another.

3. While $E[x_i y_i|\theta_i] = E[r_{xi}r_{yi}|\theta_i]$ can vary over i and need not be zero, **the unconditional covariance of x_i and y_i is zero**: $E[x_i y_i] = E[r_{xi}r_{yi}] = 0$;

The conditional variance-covariance matrix of $\{x_{it}, y_{it}\}$ is a function of $\theta_i = (\sigma_{xi}, \sigma_{yi}, r_{xi}, r_{yi})$ and has the following three elements, $\Sigma(\theta_i) \equiv \left[\sigma_{xi}^2, \sigma_{yi}^2, c_i\right]'$:

$$\Sigma(\theta_i) \equiv \begin{bmatrix} \sigma_{xi}^2 \\ \sigma_{yi}^2 \\ c_i \end{bmatrix} \,\tilde{}\, \left(\bar{\Sigma}, \Omega\right) \equiv \left(\begin{array}{c} \bar{\sigma}_x^2 \\ \bar{\sigma}_y^2 \\ \bar{c} \end{array}, \begin{array}{ccc} \omega_{xx} & & \\ \omega_{xy} & \omega_{yy} & \\ \omega_{xc} & \omega_{yc} & \omega_{cc} \end{array} \right)$$

where the conditional covariance of x_i and y_i is $c_i \equiv r_{xi}r_{yi}\sigma_{xi}\sigma_{yi}$. In specifying a joint distribution for θ_i, I also obtain a distribution for Σ_i. The mean of $\Sigma(\theta_i)$ is denoted $\bar{\Sigma} \equiv \left(\bar{\sigma}_x^2, \bar{\sigma}_y^2, \bar{c}\right)'$ and the variance-covariance matrix of $\Sigma(\theta_i)$ is denoted by Ω. The elements of Ω are labelled as follows. ω_{xy} is the covariance of σ_{xi}^2 and σ_{yi}^2, so that $\omega_{xy} > 0$ means that observations where x_i has a high variance are also observations where y_i tends to have a high variance We will refer to $\omega_{xy} \neq 0$ as "correlated

[1] $E[e_{xi}] = E[e_{yi}] = E[e_{ci}] = 0$. Differences in outcomes across observations reflect realized shocks and not differences in means. Otherwise, correlated means $cov(E[x_{it}], E[y_{it}])$ would be misidentified as correlated shocks $E[r_{xi}r_{yi}]$. This assumption could be seen merely as an interpretation of results to follow. To the econometrician, correlated expected values are observationally equivalent to correlated shocks.

I impose the additional requirement that kurtosis of e_{xi}, e_{yi}, and e_{ci}, here the fourth moment, are finite and the same for all i, so that $E\left[e_{xi}^4\right] = \kappa_x$, $E\left[e_{yi}^4\right] = \kappa_y$; $E\left[e_{ci}^4\right] = \kappa_c$.

variances." Similarly, ω_{cc} is the variance of covariance parameters c_i, so that $\omega_{cc} > 0$ will be referred to as "covariance heterogeneity." All the elements of μ and Ω are assumed to be finite. The support of σ_{xi} and σ_{yi} is positive and the support of r_{xi} and r_{xi} is $[-1, 1]$. No further restrictions are required for internal consistency. This paper tests hypotheses and imposes bounds on elements of Ω, particularly ω_{xy} and ω_{cc}.

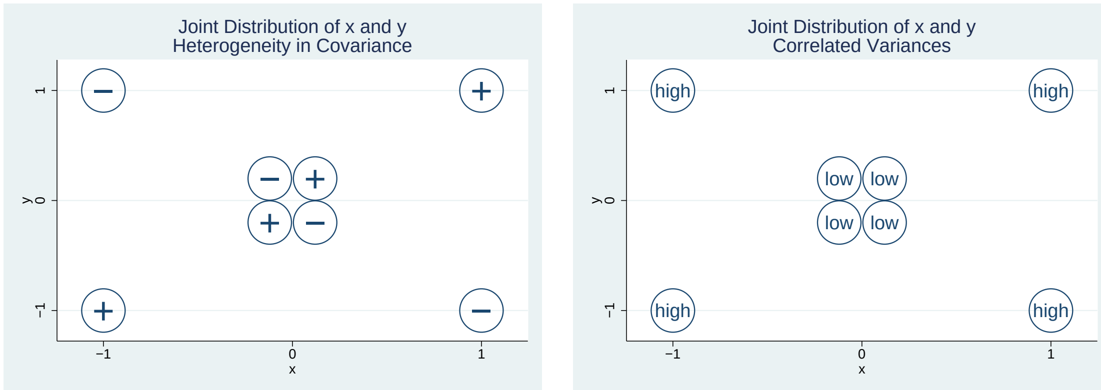
1.2 Intuition

Consider the null hypothesis of no "correlated heterogeneity," the joint hypothesis that $cov\left(\sigma_{xi}^2, \sigma_{yi}^2\right) \equiv \omega_{xy} = 0$ and $var\left(c_i\right) \equiv \omega_{cc} = 0$. This could be rejected if the sample mean of $x_i^2 y_i^2$ that is larger than it would be under independence ($\bar{\sigma}_x^2 \bar{\sigma}_y^2$). By the definitions of variance and covariance, this would imply that the sample variance of $x_i y_i$ is larger than it would be under independence ($\bar{\sigma}_x^2 \bar{\sigma}_y^2 - \bar{c}^2$), or equivalently that covariance between x_i^2 and y_i^2 is positive. In other words, correlated heterogeneity means that large absolute values of x_i will tend to coincide with large absolute values of y_i, or that x and y are not independent because they have correlated second moments. Given the additive structure of shocks proposed in Section 1.1, correlated heterogeneity indicates some combination of covariance heterogeneity, ω_{cc}, and correlated variances, ω_{xy}. To see this intuitively, note that the naive sample analogs of ω_{cc} (the sample variance of $x_i y_i$) and ω_{xy} (the sample covariance of x_i^2 and y_i^2), are both identified from the sample mean of $x_i^2 y_i^2$. Because both parameters are expressed in the same moment, it will be impossible to differentiate $\omega_{cc} > 0$ from $\omega_{xy} > 0$ without additional information.

Consider the following stylized example, where x_i and y_i both take on values of -1, 0, and 1 with probabilities $(1/4, 1/2, 1/4)$ and therefore both have a mean of 0 and variance of $\bar{\sigma}_x^2 = \bar{\sigma}_y^2 = 1/2$. When x_i and y_i are independent, the mean of $x_i^2 y_i^2$ is $1/4$ (or equivalently, the variance of $x_i y_i$ is $1/4$ or the covariance of x_i^2 and y_i^2 is 0). Now consider the joint distribution of x_i and y_i in either panel of Figure 1. Here, x_i and y_i are not independent (though they are unconditionally uncorrelated, $\bar{c} = 0$) but the marginal distributions of x_i and y_i are unchanged. The key feature of this distribution is the absence (compared with the distribution under independence) of mass where exactly one variable (x_i or y_i, but not both) is zero. Since non-zero values of x_i and y_i always coincide, the mean of $x_i^2 y_i^2$ is $1/2$ compared to $1/4$ in the case of independence (or equivalently, the variance of $x_i y_i$ is $1/2$ compared with $1/4$, and the covariance of x_i^2 and y_i^2 is $1/4$ compared to 0). Section 2 presents the test statistics used to test the null of independence of x_i and y_i when they are uncorrelated by looking at the sample mean of $x_i^2 y_i^2$ (or again, the sample variance of $x_i y_i$ or the sample covariance of x_i^2 and y_i^2). Such tests for independence are far from novel in the statistics literature.

The novelty in this paper is to decompose this rejection of independence of uncorrelated variables x_i and y_i into two types of correlated heterogeneity: correlated variances ($\omega_{xy} > 0$) and covariance heterogeneity ($\omega_{cc} > 0$). These are depicted in the two stylized depictions of hypothetical data in Figure 1. In the two panels, the unconditional joint distribution of x_i and y_i are the same; as a result, observations

Figure 1: Stylized Joint Distribution of Couples' Changes in Income



Stylized depiction of the data. See text for details.

---

have the same unconditional variance ($\bar{\sigma}_x^2 = \bar{\sigma}_y^2 = 1/2$) and covariance ($\bar{c} = 0$). The two panels differ in whether correlated variances ($\omega_{xy} > 0$) or covariance heterogeneity ($\omega_{cc} > 0$) generates the unconditional joint distribution of x and y; they differ in which parameter – σ_i^2 or c_i – varies across observations. The left panel of figure 1 shows how the excess variance in $x_i y_i$ could be explained by covariance heterogeneity, $\omega_{cc} > 0$; the right panel shows how this excess variance in $x_i y_i$ could by explained by correlated variances, $\omega_{xy} > 0$.

In the left panel, observations are either in a negative covariance state or a positive covariance state. Covariances are either -1 (observations identified with a negative sign and running from top-left to bottom-right) or 1 (observations identified with a positive sign and running from bottom-left to top-right). Conditional on c_i, the distribution is trinomial (values of -1, 0, and 1 are possible). In the framework from equation 1, $\theta_i \equiv (\sigma_{xi}, \sigma_{yi}, r_{xi}, r_{yi}) \in \{(1,1,1,1), (1,1,1,-1)\}$ with equal probability, and $e_{ci} \in \{-1, 1\}$ with equal probability.

In the right panel, some observations are in a high variance state while others are in a low variance state. Variances are either 0 (marked with a "low" and clustered at zero) or 1 (marked with a "high" and found at the corners) for both x_i and y_i and the variances for x_i and y_i are perfectly correlated. Conditional on the variances, the distribution is binomial (values of -1 and 1 are possible in the high variance state while only values of 0 or 0 are possible in the low variance state). In the framework from equation 1, $\theta_i \equiv (\sigma_{xi}, \sigma_{yi}, r_{xi}, r_{yi}) \in \{(1,1,0,0), (0,0,0,0)\}$ with equal probability, and $e_{xi}, e_{yi} \in \{-1, 0, 1\}$ with probabilities of $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{1}{4}$.

If we observe the unconditional distribution depicted in these panels, with an excess of mass in the middle and the corners so that the variance of $x_i y_i$ is larger than it would be under independence, given the framework in equation 1 we know that at least one type (or possible both) of correlated heterogeneity is present. However, it is impossible to differentiate these two possible explanations without additional data. Section 3 shows how to use panel data to place bounds on relative importance of the two types of heterogeneity.

Of course, it is possible to write down other non-additive models of the data that can generate the same distribution without parameter heterogeneity, relabelling stochastic parameters θ_i as shocks.[2] But within a simple additive model, it is possible to identify correlated parameter heterogeneity governing a pair of random variables when it is not possible to identify heterogeneity in parameters affecting the marginal distribution of a sigle variable. Compare this two-variable stylized example to the one-variable case by examining either variable in this example separately. In the left panel, the variance parameter of each individual is constant (no heterogeneity). In the right pane, the variance parameter alternates between 0 and 1 (heterogeneity). But in the right panel, shocks are conditionally binomial (and not trinomial as in the left panel) so kurtosis is lower. Since the marginal distributions are the same in either panel, there is no way to differentiate these possibilities. In the two-variable case, there are again two possibilities that cannot be differentiated, but both involve some type of parameter heterogeneity.

2 Testing for Correlated Heterogeneity

Here, I present the distribution for a test statistic based on the covariance of second moments under the null independence (no "correlated heterogeneity"). Given the framework in Section 1.1, this is the joint hypothesis that $cov\left(\sigma_{xi}^2, \sigma_{yi}^2\right) \equiv \omega_{xy} = 0$ and $var\left(c_i\right) \equiv \omega_{cc} = 0$. Coupled with the assumption 5 that shocks are unconditionally uncorrelated $E\left[x_i y_i\right] \equiv \bar{c} = 0$, these imply that shocks are conditionally uncorrelated as well. Therefore, $E\left[x_i y_i | \theta_i\right] \equiv c_i = c_i = 0$ for each i and therefore $r_{xi} = 0$, $r_{yi} = 0$, $\sigma_{xi} = 0$, or $\sigma_{yi} = 0$ (or several of these) for each i.

Before assuming the null, conditional expectations can be identified by substitution from (1) to get equation (2):

$$E\left[x_i^2 y_i^2 | \theta_i\right] = \sigma_{xi}^2 \sigma_{yi}^2 + c_i^2\left(\kappa_c - 1\right) \tag{2}$$

which by the definitions of variance and covariance implies that $var\left(x_i y_i | \theta_i\right) = \sigma_{xi}^2 \sigma_{yi}^2 + c_i^2\left(\kappa_c - 2\right)$ and $cov\left(x_i^2, y_i^2 | \theta_i\right) = c_i^2\left(\kappa_c - 1\right)$. Note the critical role of the first assumption from Section 1.1 (additive shocks) in reaching the parsimonious, additive structure of equation (2). Integrating $E\left[x_i^2 y_i^2 | \theta_i\right]$ over θ_i to obtain unconditional expectations yields equation (3):

$$\begin{aligned} E\left[x_i^2 y_i^2\right] &= \bar{\sigma}_x^2 \bar{\sigma}_y^2 + \omega_{xy} + \left(\omega_{cc} + \bar{c}^2\right)\left(\kappa_c - 1\right); \\ var\left(x_i y_i\right) &= \bar{\sigma}_x^2 \bar{\sigma}_y^2 + \left(\kappa_c - 2\right)\bar{c}^2 + \omega_{xy} + \omega_{cc}\left(\kappa_c - 1\right) \end{aligned} \tag{3}$$

which again by the definitions of covariance implies that $cov\left(x_i^2, y_i^2\right) = \omega_{xy} + \left(\omega_{cc} + \bar{c}^2\right)\left(\kappa_c - 1\right)$.

[2] For example, the left panel can be relabelled with the following model: $x_i = e_{ci} e_{xi}; y_i = e_{ci}$ where $e_{ci} \in \{-1, 0, 1\}$ with probabilities of $\frac{1}{4}$, $\frac{1}{2}$, and $\frac{1}{4}$ and $e_{xi} \in \{-1, 0, 1\}$ with equal probabilities. The right panel can be relabelled with the following model: $x_i = e_{xi} e_{ci}; y_i = e_{yi} e_{ci}$ where $e_{xi}, e_{yi} \in \{-1, 1\}$ with equal probability and $e_{ci} \in \{0, 1\}$ with equal probability. In each case shocks which enter multiplicatively take the place of stochastic parameters.

Note that $\bar{\sigma}_x^2$, $\bar{\sigma}_y^2$, and $\bar{c}$ can be estimated from the data with $\frac{1}{N}\Sigma x_i^2$, $\frac{1}{N}\Sigma y_i^2$, and $\frac{1}{N}\Sigma x_i y_i$, respectively. Under the null hypothesis that $\omega_{xy} = 0$ and $\omega_{cc} = 0$, equation (3) reduces to $E\left[x_i^2 y_i^2\right] = \bar{\sigma}_x^2 \bar{\sigma}_y^2 + \bar{c}^2 (\kappa_c - 1)$ which implies that $var\left(x_i y_i\right) = \bar{\sigma}_x^2 \bar{\sigma}_y^2 + \bar{c}^2 (\kappa_c - 2)$ and $cov\left(x_i^2, y_i^2\right) = \bar{c}^2 (\kappa_c - 1)$.

I consider settings where x_i and y_i are unconditionally uncorrelated, so that the covariance is approximately zero on average, so that $\bar{c} \approx 0$. This is the second major assumption in Section 1.1. While obviously an extremely strong assumption, it is sometimes an empirically relevant one. For example, the average covariance of couples' income changes has been estimated at close to zero.Under the null hypothesis that $\omega_{xy} = \omega_{cc} = 0$, $E\left[x_i^2 y_i^2\right] = var\left(x_i y_i\right) = \bar{\sigma}_x^2 \bar{\sigma}_y^2$ and $cov\left(x_i^2, y_i^2\right) = 0$.

Under the null hypothesis when $\bar{c} = 0$, for a randomly chosen i from the population, $x_i^2 y_i^2$ will be drawn from the following distribution,

$$x_i^2 y_i^2 \sim \left(\bar{\sigma}_x^2 \bar{\sigma}_y^2, \left(\omega_{xx} + \bar{\sigma}_x^4\right) \kappa_x \left(\omega_{yy} + \bar{\sigma}_y^4\right) \kappa_y - \bar{\sigma}_x^4 \bar{\sigma}_y^4\right). \tag{4}$$

When $\bar{c}$ is not zero, the mean of the distribution from which $x_i^2 y_i^2$ is drawn will then be $\bar{\sigma}_x^2 \bar{\sigma}_y^2 + \bar{c}^2 (\kappa_c - 2)$. [3]

When $\bar{c}$ is small relative to $\bar{\sigma}_x \bar{\sigma}_y$, $\bar{c}$ will be unimportant to the mean of the distribution unless tails are so fat that kurtosis is nearly infinite. For example, in the empirical example presented in this paper, $\bar{c} \approx -0.07 \bullet \bar{\sigma}_x \bar{\sigma}_y$ so that $\bar{\sigma}_x^2 \bar{\sigma}_y^2 + \bar{c}^2 (\kappa_c - 2) \approx (1 + (\kappa_c - 2) \bullet 0.0049) \bullet \bar{\sigma}_x^2 \bar{\sigma}_y^2$. Given such a small value for $\bar{c}$, κ_c would have to be implausibly large for this additional term to have a substantive effect on the mean of the distribution. Note that under the null hypothesis, $\left(\omega_{xx} + \bar{\sigma}_x^4\right) \kappa_x$ can be estimated with $\frac{1}{N}\Sigma_i x_i^4$ and $\left(\omega_{yy} + \bar{\sigma}_y^4\right) \kappa_y$ can be estimated with $\frac{1}{N}\Sigma y_i^4$. Even when $\bar{c} \neq 0$ so that the variance term is complex, it can still be estimated in sample with $\frac{1}{N}\Sigma_i x_i^4 y_i^4 - \left(\frac{1}{N}\Sigma_i x_i^2 y_i^2\right)^2$ so long as higher-order moments are finite. As with the mean, small deviations from $\bar{c} = 0$ will have little impact on the variance of the distribution. Since observations are assumed to be *iid*, under the null hypothesis with $\bar{c} = 0$ the sample variance, $\frac{1}{N}\Sigma_i x_i^2 y_i^2 - \left(\frac{1}{N}\Sigma_i x_i y_i\right)^2$, has the following distribution:

$$\frac{1}{N}\Sigma_i x_i^2 y_i^2 \sim \left(\left(\frac{1}{N}\Sigma_i x_i y_i\right)^2, \frac{1}{N}\left(\left(\omega_{xx} + \bar{\sigma}_x^4\right) \kappa_x \left(\omega_{yy} + \bar{\sigma}_y^4\right) \kappa_y - \bar{\sigma}_x^4 \bar{\sigma}_y^4\right)\right) \tag{5}$$

Recall that each of the terms is finite and can be estimated from sample data. Since we have the distribution of the sample variance it is straightforward to test that null. The key thing to note is that κ_c (which cannot be observed without assuming the null) does not show up in the test when $\bar{c} = 0$ and therefore does not need to be estimated. As long as $\bar{c}$ is relatively close to zero, κ_c will be unimportant.

Formally, the sample moment $\frac{1}{N}\Sigma_i x_i^2 y_i^2$ just allows for a test of the independence of shocks, $E\left[f\left(x_i\right) f\left(y_i\right)\right] = E\left[f\left(x_i\right)\right] E\left[f\left(y_i\right)\right]$. Independence requires that this be true for all $f\left(\right)$ and $g\left(\right)$ and here we look only at second moments, $f\left(x_i\right) = x_i^2$

[3] The variance of the distribution will take a more complex and uninformative form. Unlike the case with $\bar{c} = 0$, when $\bar{c}$ is not zero the variance will have elements with moments above the fourth moment such as $E\left[e_{xi}^4\right]$, so that possibility of infinite higher-order moments would have to be ruled out. Truncating the data would assure this.

and $g(y_i) = y_i^2$. The novelty here is that equation 3 decomposes this rejection of independence into two types of correlated heterogeneity. There is empirically relevant because there are variety of settings where covariance heterogeneity is interesting and where estimates of the mean covariance are close to zero. For example, do couples' incomes all jointly evolve in the same way?

2.1 "Wife-Swap Bootstrap" Test

So far, we have assumed that observations are independent. For a cross-section of randomly chosen individuals who face idiosyncratic shocks, this assumption may be relatively innocuous. When data comes from a panel, this is seldom true. For notation, I place time subscripts on data (x_{it} and y_{it}), shocks (e_{xit},e_{yit},e_{cit}), parameters (θ_{it}) and hyperparameters ($\bar{\Sigma}_t$, Ω_t). In this case, the sample variance, $\frac{1}{N}\frac{1}{T}\Sigma_i\Sigma_t x_{it}^2 y_{it}^2 - \left(\frac{1}{N}\frac{1}{T}\Sigma_i\Sigma_t x_{it} y_{it}\right)^2$ will be drawn from a distribution with same mean as in the *iid* case, $\bar{\sigma}_x^2\bar{\sigma}_y^2$, but not the same variance:

$$\frac{1}{N}\frac{1}{T}\Sigma_i\Sigma_t x_{it}^2 y_{it}^2 - \left(\frac{1}{N}\frac{1}{T}\Sigma_i\Sigma_t x_{it} y_{it}\right)^2 \sim \left(\bar{\sigma}_x^2\bar{\sigma}_y^2, \frac{1}{NT}\left(\begin{array}{c} \left(\omega_{xx}+\bar{\sigma}_x^4\right)\kappa_x\left(\omega_{yy}+\bar{\sigma}_y^4\right)\kappa_y - \bar{\sigma}_x^4\bar{\sigma}_y^4 \\ +\frac{1}{T}\Sigma_i\Sigma_t\Sigma_{s\neq t} cov\left(x_{is}^2 y_{is}^2, x_{it}^2 y_{it}^2\right) \end{array}\right)\right) \quad (6)$$

While the first part of the variance is trivial to estimate from sample data as $\frac{1}{NT}\left(\frac{1}{N^2T^2}\Sigma_i\Sigma_t x_{it}^4 \Sigma_i\Sigma_t y_{it}^4\right.$ the covariance terms are more difficult to estimate. The main challenge in a non-rectangular panel is that attrition may be related to the autocorrelation. Without attrition, $cov\left(x_{is}^2 y_{is}^2, x_{it}^2 y_{it}^2\right)$ can be estimated from data under the null as $\frac{1}{N}\Sigma_i x_{is}^2 x_{it}^2 \frac{1}{N}\Sigma_i y_{is}^2 y_{it}^2$. An alternative way to obtain the same variance can be obtained by noting that under the null, $var\left(\frac{1}{N}\frac{1}{T}\Sigma_i\Sigma_t x_{it}^2 y_{it}^2\right) = var\left(\frac{1}{N}\frac{1}{T}\Sigma_i\Sigma_t x_{it}^2 y_{jt}^2\right)$ for a randomly chosen $j \neq i$. (The non-rectangularity problem can be overcome if j is chosen so that i and j have the same number of observations.) As a result, it is straightforward obtain the variance of the estimator by repeatedly sampling $\frac{1}{N}\frac{1}{T}\Sigma_i\Sigma_t x_{it}^2 y_{jt}^2$ for different choices of j and taking the variance of these. When x and y refer to the incomes of husbands and wives, this involves randomly pairing all husbands and wives from the data, and calculting the estimator for this synthetic pair. Doing this repeatedly builds up a reference distribution under the null. I use the tongue-in-cheek name *wife-swap bootstrap* to refer to this procedure.

Note that we are not taking advantage of panel features of the data, but rather are identifying the loss of statistical power (and therefore increased confidence intervals) that come with correlated data.

3 Bounding Correlated Heterogeneity

Once we reject the null of no correlated heterogeneity, we know that either $\omega_{xy} \neq 0$ or $\omega_{cc} \neq 0$ or both. $\omega_{cc} > 0$ indicates that the covariance differs across observations. $\omega_{xy} > 0$ indicates that observations with high-variance x also tend to have high-variance y. While ω_{cc} and ω_{xy} are identified from the same moment in the data

$\frac{1}{N}\Sigma_i x_i^2 y_i^2$, they reflect completely different phenomenon Consider the application to couples, and x refers to the change in husbands' incomes and y refers to the change in wives' incomes. $\omega_{cc} > 0$ indicates that some couples incomes co-move more closely than others. This could be interpreted as saying that the diversification benefits of marriage vary across couples. Identifying this heterogeneity, $\omega_{cc} > 0$, is needed if one is to correct the attenuation bias caused when a noisy estimate of a couples' covariance, $x_i y_i$, is used as a right-hand-side variable in a regression to predict divorce or consumption. By contrast, ω_{xy} identifies assortative mating in risk, whether husbands with volatile incomes have wives with volatile incomes. This sort of assortative mating could be a test for theories of optimal partner selection.

Most simply, if the sample variance is smaller than it would be under independence, then we know that $\omega_{xy} < 0$. From equation 3, when $E\left[x_i^2 y_i^2\right] < \sigma_x^2 \sigma_y^2$, then $\omega_{xy} < 0$ since $\kappa \geq 1$ (the lowest possible kurtosis is 1, which is the value for a binomial distribution) and $\omega_{cc} \geq 0$ (no variance can be negative), and $\bar{c} = 0$ (by assumption). When the sample variance is larger than it would be under independence, additional information from other covariates of panel data is needed.

3.1 Panel Data

To separate, $\omega_{xy} > 0$ from $\omega_{cc} > 0$, I exploit minimal features of panel data. I assume that there exist s and t sufficiently far apart (for example, a fixed distance k) that common shocks from the two periods are uncorrelated, $E\left[e_{xis} e_{xit}\right] = E\left[e_{yis} e_{yit}\right] = E\left[e_{cis} e_{cit}\right] = 0$ and therefore $E\left[x_{is} y_{it}\right] = 0$, but close enough that distributions are approximately stable between the period $\bar{\Sigma}_s = \bar{\Sigma}_t$, $\Omega_s = \Omega_t$. Note that the assumption of stability does not mean that c_{it} remains constant over time, only that the distribution of c_{it} over all i remains constant over time. These assumptions are strong, but are readily testable in the case of couples, whose non-overlapping income changes are nearly uncorrelated and where any changes in the distribution of parameters is slow. For example, Abowd and Card (1989) show that innovations to income are not autocorrelated at lags greater than two years. We now rewrite equation (3) with time subscripts:

$$E\left[x_{it}^2 y_{it}^2\right] = \omega_{xyt} + \bar{\sigma}_{xt}^2 \bar{\sigma}_{yt}^2 + \left(\omega_{cct} + \bar{c}_t^2\right)\left(\kappa_{ct} - 1\right). \tag{7}$$

Our goal is to use information about $E\left[x_{it}^2 y_{it}^2\right]$, $\bar{\sigma}_{xt}^2 \bar{\sigma}_{yt}^2$, and $\bar{c}_t^2$ (which can be estimated from sample data) to place bounds on ω_{xyt}, ω_{cct}, and κ_{ct}.

ω_{cc} : Most obviously, note that $\left|cov\left(c_{is}, c_{it}\right)\right| < \sqrt{var\left(c_{is}\right) var\left(c_{it}\right)} = var\left(c_{it}\right) \equiv \omega_{cc}$ (the last equality by the stability assumption). $cov\left(c_{is}, c_{it}\right)$ can be readily estimated from the data as $\frac{1}{NT}\Sigma_t \Sigma_i \left(x_{is} y_{is} x_{it} y_{it} - \bar{c}^2\right)$. This estimates a lower bound for ω_{cc}. The slower c_{it} evolves the tighter this lower bound will be.

ω_{xy} : While it is not strictly required by the assumptions above, all but the most pathological distributions will exhibit

$$\frac{1}{2}\left|cov\left(\sigma_{xis}^2, \sigma_{yit}^2\right) + cov\left(\sigma_{yis}^2, \sigma_{xit}^2\right)\right| < \frac{1}{2}\left(cov\left(\sigma_{xis}^2, \sigma_{yis}^2\right) + cov\left(\sigma_{xit}^2, \sigma_{yit}^2\right)\right) = \omega_{xy}$$

where the last equality follows from the assumption of stability. Contemporaneous shocks should be more highly correlated than lead or lagged shocks with a large enough time-gap. This need not be true when one variable predicts subsequent values for other, but when $cov\left(\sigma^2_{xis}, \sigma^2_{yit}\right)$ and $cov\left(\sigma^2_{yis}, \sigma^2_{xit}\right)$ are positive and similar in value, contemporaneous shocks are more likely to have similar magnitudes. $cov\left(\sigma^2_{xis}, \sigma^2_{yit}\right)$ and $cov\left(\sigma^2_{yis}, \sigma^2_{xit}\right)$ can be readily estimated from the data with $\frac{1}{NT}\Sigma_t\Sigma_i\left(x^2_{is}y^2_{it} - \bar{\sigma}^2_x\bar{\sigma}^2_y\right)$ and $\frac{1}{NT}\Sigma_t\Sigma_i\left(x^2_{it}y^2_{is} - \bar{\sigma}^2_x\bar{\sigma}^2_y\right)$, respectively. This estimates a lower bound $\underline{\omega}_{xy}$ on ω_{xy} as $\frac{1}{2}\left(\frac{1}{NT}\Sigma_t\Sigma_i\left(x^2_{is}y^2_{it} - \bar{\sigma}^2_x\bar{\sigma}^2_y\right) + \frac{1}{NT}\Sigma_t\Sigma_i\left(x^2_{it}y^2_{is} - \bar{\sigma}^2_x\bar{\sigma}^2_y\right)\right)$. Note that lower bounds on both ω_{cc} and ω_{xy} have a single-variable analog. One can obtain a lower bound on heterogeneity in x, ω_{xx}, with $cov\left(\sigma^2_{xis}, \sigma^2_{xit}\right)$.

If one is willing to assume a degree of kurtosis, κ_c, a lower bound $\underline{\omega}_{xy}$ on ω_{xy} places an upper bound $\bar{\omega}_{cc}$ on ω_{cc} and *vise versa*. In particular, plugging the lower bound $\underline{\omega}_{xy}$ into equation 7 yields:

$$\bar{\omega}_{cc} = \left(E\left[x^2_{it}y^2_{it}\right] - \underline{\omega}_{xy} - \bar{\sigma}^2_x\bar{\sigma}^2_y\right) / \left(\kappa_c - 1\right) - \bar{c}^2$$

While the assumption of κ_c is arbitrary here, results are much less sensitive to misspecification than in the analogous case for variance heterogeneity affecting the marginal distribution of a single variable. When examining heterogeneity in variance, ω_{xx}, it is straightforward to rewrite equation 7 when $x_{it} = y_{it}$, in which case $c_{it} = \sigma^2_{xit}$, $\bar{c}_t = \bar{\sigma}^2_{xt} = \bar{\sigma}^2_{yt}$, and $\omega_{xxt} = \omega_{xyt} = \omega_{cct}$:

$$\begin{aligned} E\left[x^4_{it}\right] &= \left(\omega_{xx} + \bar{\sigma}^4_x\right)\kappa_x \\ \omega_{xx} &= E\left[x^4_{it}\right]/\kappa_x - \bar{\sigma}^4_x \end{aligned}$$

To examine the impact of mis-specification of κ_t on estimates of ω_{xxt}, consider the elasticity of ω_{xxt} with respect to changes in κ_t:

$$\frac{d\omega_{xx}}{d\kappa_x}\frac{\kappa_x}{\omega_{xx}} = -\frac{\omega_{xx} + \bar{\sigma}^4_x}{\omega_{xx}}.$$

When heterogeneity in variance (ω_{xxt}) is small compared to variance ($\bar{\sigma}^4_{xt}$), then small misspecification will lead to enormous mis-specification of ω_{xxt}. By contrast, small changes in κ_{ct} have a more modest impact on estimation of the upper bound on ω_{xyt}. Rearranging equation 7 and differentiating yeilds:

$$\frac{d\omega_{cc}}{d\kappa_c}\frac{\kappa_c}{\omega_{cc}} = -\frac{\kappa_c}{\kappa_c - 1}.$$

Unless kurtosis is implausibly small (<3, where recall that 3 is the kurtosis for a normal distribution and 1 the kurtosis for a binomial distribution), this elasticity will be relatively close to one. As a result, small mis-specification of kurtosis (κ_c) will have a much smaller impact on estimates of covariance heterogeneity (ω_{cc}).

κ_c : The lower bounds $\underline{\omega}_{cc}$ and $\underline{\omega}_{xy}$ placed on ω_{cc} and ω_{xy}, respectively, imply an upper bound $\bar{\kappa}_c$ on κ_c as:

$$\bar{\kappa}_c = 1 + \left(E\left[x_{it}^2 y_{it}^2\right] - \underline{\omega}_{xy} - \bar{\sigma}_x^2 \bar{\sigma}_y^2\right) / \left(\underline{\omega}_{cc} + \bar{c}^2\right).$$

3.2 Additional Covariates

Additional covariates can also be used to identify correlated heterogeneity. Consider the regression:

$$x_{it} y_{it} = \alpha + \beta_c' Z_{it} + e_{cit}$$

Here Z_{it} is the set of covariates. Note that $\omega_{cc} > var\left(\beta_c' Z_{it}\right) = \beta_c' Z_{it} Z_{it}' \beta_c$. If we can identify variables Z_{it} that explain some of the variation in $x_{it} y_{it}$, this places a lower bound on the variation in covariance parameters, $\underline{\omega}_{cc}$. As with the panel approach considered in Section 3.1, knowing $\underline{\omega}_{cc}$ and κ_c gives an upper bound on the degree to which variances are correlated, $\bar{\omega}_{xy}$.

Similarly, consider the regressions:

$$\begin{aligned} x_{it}^2 &= \alpha + \beta_x' Z_{it} + e_{xit}; \\ y_{it}^2 &= \alpha + \beta_y' Z_{it} + e_{yit} \end{aligned}$$

Some of the correlated variances can be explained with $cov\left(\beta_x Z_{it}, \beta_y Z_{it}\right) = \beta_x' Z_{it} Z_{it}' \beta_y'$. The problem here is that $E\left[e_{xit} e_{yit}\right]$ is unknown, so the amount of correlated variances explained by covariates $\left(\beta_x' Z_{it} Z_{it}' \beta_y'\right)$ sheds little light on overall amount of correlated variance unless a large fraction of variation can be explained with it, so that $\beta_x' Z_{it} Z_{it}' \beta_y'$ is large compared with $E\left[e_{xit}^2\right] E\left[e_{yit}^2\right]$.

4 An Application to Couples' Incomes

Here, I show these methods as applied to the joint income process for couples. Husbands' and wives' income changes are roughly unconditionally uncorrelated. However, they do not move independently, as their squared income changes are correlated. Do some couples' incomes move together while other couples' incomes move in opposite directions? When the variance of a husband's income shock is large, does the same tend to be true for his wife? When we expect that the co-movement of couples' incomes may not be constant over the life of a marriage, the tools developed here are necessary to answer these questions. Furthermore, these questions are of practical interest if we aim to use a noisy estimate of c_{it} (namely $x_{it} y_{it}$) as a right-hand-side variable to predict outcomes of interest such as household saving or consumption as in Shore (2007). Coefficients in this regression will have substantial attenuation bias that can be corrected with a reliability ratio, $\omega_{cc}/var\left(x_{it} y_{it}\right)$ (or a reliability matrix when c_{it}, x_{it} and/or y_{it} are autocorrelated).

4.1 Data

The data and estimation procedure outlined here follow Shore (2006) exactly. While this paper is intended to be self-contained in outlining data, methods, and assumptions, it omits the lengthly data description found in that paper. Interested readers may wish consult that paper to learn more about the features of this data.

Data are drawn from the Panel Study of Income Dynamics (PSID). The PSID is a nationally representative panel of U.S. households that has tracked families annually from 1968 to the present. Data are not collected in even-numbered years after 1997; this paper uses data collected through 2005. However, since most analyses use one-year income changes, only data through 1997 will be used in most circumstances. The PSID includes data on households, including household food consumption and the education, income, hours worked, employment status, and age of husbands and wives. I use annual labor income as a measure of income. I restrict the sample to married couples, to couples where the marriage is the husband's first, to observations for which both the husband and wife are between the ages of 22 and 60, and for which the couple has been married for no more than 35 years.

I remove the predictable (to the econometrician) component of income and examine the time series properties of the unpredictable component, excess log income. As is common in the literature, this excess log income is the residual from a least-squares regression of the natural log of labor income (for either the husband or the wife) on the following regressors: a cubic in age for each level of educational attainment (none, elementary, junior high, some high school, high school, some college, college, graduate school) for both husband and wife, a cubic in the number of years the couple has been married, the presence and number of infants, young children, and older children in the household, the total number of family members in the household, and dummy variables for each calendar year. So that log income results are not dominated by income values close to zero, I limit the regression sample to individuals who earn at least $\$1,000$ (in 2001 dollars).

The residuals from this regression are Winsorized at the 5^{th} and 95^{th} percentiles, so that residuals below the 5^{th} percentile are replaced by the 5^{th} percentile value and those above the 95^{th} percentile are replaced by the 95^{th} percentile value. At the same time, values omitted from the initial regression because real annual income was below $\$1,000$ are given the 5^{th} percentile residual value. The vast majority of these initially omitted values have an income of exactly zero. This reduces selection bias by including extreme values, while at the same time limiting the degree to which such outlier drive the results. Even more important, it allows us to exploit variation coming from transitions into and out of the labor force. One-year changes are demeaned.

Table 1 presents summary statistics on one-year changes in excess log income. Note that most one-year income changes are relatively small. Since these changes are in logs, small changes approximate percentage changes. The inter-quartile ranges for wives (x_{it} from -10 percent to 8 percent) and husbands (y_{it} from -8 percent to 10 percent) are modest. However, there are occasional very large changes in income, so that the standard deviations of one-year income changes (55 percent and 32 percent,

Table 1: Distribution of Spouses' Change in Excess Log Income

Variable	x_{it}	y_{it}
Spouse	Wife	Husband
Mean	0	0
St. Dev.	0.5490	0.3184
Observations	20,762	20,762
Minimum	-2.8499	-1.8283
5^{th} Percentile	-0.8390	-0.5258
25^{th} Percentile	-0.0955	-0.0796
50^{th} Percentile	-0.0179	0.0064
75^{th} Percentile	0.0806	0.0987
95^{th} Percentile	0.9305	0.4708
Maximum	2.8141	1.8410

This table presents the distributions of one-year changes in Winsorized excess log income for wives and hubands, x_{it} and y_{it}, respectively. The construction of Winsorized excess log incomes is explained in the text. In brief, annual log labor incomes for husbands and wives are separately regressed on a host of covariates. The residuals from these regressions are Winsorized at the 5^{th} and 95^{th} percentiles. These changes are de-meaned, so means are zero by construction. The median one-year change would be exactly zero in the absence of de-meaning, so -1 times the median values gives the average annual change. The sample is limited to observations where data exists in the six years prior to the year in question.

respectively) are much larger than the interquartile ranges. These fat-tails could be the result of fat-tailed shocks (occasional large income changes) or heterogeneity (some observations are expected to have larger variances while others are expected to have smaller variances, though conditional on these variances tails are not fat).

The patterns of autocorrelation are presented in Table 2. The left-panel shows the correlations of wives' (x_{it}, left column) and husbands' (y_{it}, right column) one-year changes in excess log income with the one- through four-year lags (1^{st} through 4^{th} row) of wives' one-year changes in excess log income. The right panel is the same, though with husbands' lag income changes. While there is strong autocorrelation at a one-year lag (increases in income tend to be followed by decreases in the following year) for both husbands and wives, autocorrelations drop off rapidly at greater lags. While small, autocorrelations at lags greater than one year are larger here than in Abowd and Card (1989), primarily because income changes are Winsorized. For this paper, the key thing to note is that one spouse's income changes are nearly uncorrelated with the lagged changes in the other's income.

Table 2: Caption: Spouses' Autocorrelations and Lead-Lag Correlations of One-Year Changes in Income

Autocorrelation for Wives

correlation	x_{it}	y_{it}
x_{it-1}	-0.2133	0.0001
x_{it-2}	-0.0766	-0.0028
x_{it-3}	-0.0251	0.0151
x_{it-4}	-0.0395	0.0094

Autocorrelation for Husbands

correlation	x_{it}	y_{it}
y_{it-1}	-0.0191	-0.3193
y_{it-2}	-0.0139	-0.0445
y_{it-3}	0.0173	-0.0217
y_{it-4}	-0.0022	-0.0169

x_{it} and y_{it} are the one-year changes in wives' and husbands' Winsorized excess log incomes. The left-panel shows the correlations of wives' (x_{it}, left column) and husbands' (y_{it}, right column) one-year changes in excess log income with the one- through four-year lags (1^{st} through 4^{th} row) of wives' one-year changes in excess log income. The right panel is the same, though with husbands' lag income changes.

4.2 Measures of Co-Movement

Here, I present a simple income process that is consistent with the structure outlined in Section 1.1. Model parameters from this process may differ across couples. Before testing for correlated heterogeneity in the Section 4.3, I must first identify the parameters that may vary across couples and the moments in the data that identify them. While more complex income processes are possible, it is standard in the literature to assume that excess log income is composed of permanent (p) and transitory (ε) components:

$$\begin{aligned} z_{yit} &= p_{yit} + \varepsilon_{yit}; \\ p_{yit} &= p_{yiT_0^i} + \sum_{\tau=T_0^i+1}^{t} \omega_{yi\tau}. \end{aligned} \tag{8}$$

Here, z_{yit} refers to the excess log income of the husband in household i in year t. The same process could be applied to wives as well, with xs replacing ys. x_{it} and y_{it} will be defined as changes in excess log income over an interval, $x_{it} \equiv z_{xit} - z_{xit-k}$ and $y_{it} \equiv z_{yit} - z_{yit-k}$. The length of these intervals will vary across the various estimators described below. In equation 8, transitory income, ε_{yit}, is assumed to be i.i.d. with variance $\sigma^2_{\varepsilon yit}$; permanent income, p_{yit}, is assumed to have a unit root so that innovations to permanent income, $p_{yit} - p_{yit-1} = \omega_{yit}$, are i.i.d. with variance $\sigma^2_{\omega yit}$. Subsequently, "transitory variance" refers to the variance of transitory income, $\sigma^2_{\varepsilon yit}$; "permanent variance" refers to the variance of innovations to permanent income, $\sigma^2_{\omega yit}$.

If husbands' and wives' incomes individually evolve as in equations 8, it is natural to consider the joint income process where couples' income shocks may be correlated. For couple i at time t, I consider $E[\omega_{xit}\omega_{yit}] \equiv c_{\omega it}$ and $E[\varepsilon_{xit}\varepsilon_{yit}] \equiv c_{\varepsilon it}$, which I subsequently refer to as the "permanent covariance" and the "transitory covariance."

While husbands' transitory shocks may be correlated with wives' permanent ones, and *vice versa*, these cross-covariances are assumed to be zero here.

In this setting, Shore (2006) proposes three simple estimators to identify various features of the covariance of couples' income changes:

1. Raw: The simplest measure of the covariance of couples' income changes is the product of contemporaneous income changes. This raw estimate of covariance is simply $\hat{\delta}_{rit} \equiv x_{rit}y_{rit} \equiv (z_{xit} - z_{xit-k})(z_{yit} - z_{yit-k})$, and captures a mix of permanent and transitory shocks over the interval from $t-k$ to k. Under the specified income process $E[x_{rit}y_{rit}|\theta_{it}] = c_{\varepsilon it} + c_{\varepsilon it-k} + \Sigma_{s=t-k+1}^{t} c_{\omega is}$. In this paper, we take $k=1$. Sample moments based on this product are presented in the first column of Table 3. The sample mean of $\hat{\delta}_{rit} = x_{rit}y_{rit}$ is close to zero, with an implied correlation of less than one percent in absolute value.

2. Permanent: To isolate the permanent covariance without contamination from the transitory variance, Shore (2006) proposes the product of a wife's short term change in income and her husband's long-term change in income that spans this short term change. This estimate of the permanent covariance is then $\hat{\delta}_{\omega it} \equiv x_{\omega it}y_{\omega it} \equiv (z_{xit} - z_{xit-k})(z_{yit+m} - z_{yit-k-m})$, and $E[x_{\omega it}y_{\omega it}|\theta_{it}] = c_{\omega it}$. An advantage of this measure is that it isolates the permanent covariance even when the income process is much more general than the one specified here. It will identify the permanent covariance when permanent shocks enter in over at most m periods (as opposed to one period in the specified income process), when transitory shocks damp out in at most m periods (as opposed to one period in the specified income process), or when cross-covariances are non-zero (as ruled out with this income process). In this paper, I use $k=1$ and (following Abowd and Card (1989)) $m=2$. Sample moments based on this product are presented in the second column of Table 3. The sample mean of $(z_{xit} - z_{xit-1})(z_{yit+2} - z_{yit-3})$ is slightly negative but close to zero, with an implied correlation of about -2 percent.

3. Transitory: Under the specified income process, the transitory covariance can be identified by looking at the product of income changes for one spouse and their lag for the other spouse, where the beginning of one interval corresponds to the end of the other. The estimator of the transitory covariance is then $-\hat{\delta}_{\varepsilon it} \equiv x_{\varepsilon it}y_{\varepsilon it} \equiv (z_{xit} - z_{xit-k})(z_{yit+k} - z_{yit})$. Under the specified income process, its expectation is $E[x_{\varepsilon it}y_{\varepsilon it}|\theta_i] = -c_{\varepsilon it}$. In this paper, I use $k=1$. Sample moments based on this product are presented in the third column of Table 3. The sample mean of $(z_{xit+1} - z_{xit})(z_{yit} - z_{yit-1})$ is slightly negative but close to zero, with an implied correlation of about -1 percent.

Note that even when the income process is mis-specified, it is still meaningful to look for heterogeneity using these estimators. When the income process is mis-specified, we are merely identifying heterogeneity in $E\left[\hat{\delta}_{rit}|\theta_i\right]$, $E\left[\hat{\delta}_{\omega it}|\theta_{it}\right]$, or $E\left[\hat{\delta}_{\varepsilon it}|\theta_i\right]$, though this will map to heterogeneity in different underlying model parameters.

Table 3 presents a variety of sample moments relevant for identifying heterogeneity in couples' income changes. Recall that estimates of the average variance $\bar{\sigma}_x^2$ and $\bar{\sigma}_y^2$ are required for the correlated heterogeneity tests outlined in this paper. Note that these will differ across estimators with the interval over which income changes are taken (in the case of y_{it}) and with the sample. Samples will differ because each estimator spans a different number of years of data, and only observations are included in a given column when each moment is defined for that observation.

4.3 Testing for Correlated Heterogeneity

If one is willing to igore the panel feature of the data (erroneously assuming that x_{it} and y_{it} are not autocorrelated), then the sample moments from Table 3 provide estimates of all of the parameters needed to test the null of no correlated heterogeneity. Table 4 presents the results of these tests for each estimator of covariance.

We can strongly reject the hypothesis of no correlated heterogeneity for the all three measures of covariance. The top panel presents the sample variances of these three products, $x_{it}y_{it}$ for three different measures of x_{it} and y_{it}as outlined in Section 4.2. The lower panels present the distribution that this sample variance would have under three null hypotheses. Each of these nulls imposes that there is no correlated heterogeneity, $\omega_{cc} = 0$ and $\omega_{xy} = 0$; they differ in the assumptions made about other features of the distribution. Results under all three nulls are similar.

Null 1 assumes that a) all observations are independent and b) husbands' and wives shocks are uncorrelated so that $\bar{c} = 0$. While a) is clearly untrue, Table 3 shows that b) is approximately true. The distribution under this null is taken from equation 5, where parameters are replaced by the sample moments from Table 3 that estimate them. Null 2 is identical to Null 1, except that the assumption that $\bar{c} = 0$ is relaxed and $\bar{c}$ is estimated from the data. Section 2 shows howt to calculate the mean and variance of the distribution under the null in this case. Here, I assume that shocks are normal so that $\kappa = 3$, though given that $\bar{c}$ is close to zero results are extremely insensitive to the choice of κ. Null 3 uses the "wife-swap bootstrap" to allow observations for a given couple to be autocorrelated. This approach again imposes that $\bar{c} = 0$, though a comparison of Nulls 1 and 2 shows that this assumption is innocuous given the data. Because shocks are negatively correlated, correcting for autocorrelation actually tightens the distribution of possible values for the sample variance under the null. Under all three nulls and for all three measures of couples' income changes, the hypothesis of no correlated heterogeneity can be strongly rejected.

4.4 Decomposing Couples' Correlated Heterogeneity

After rejecting the joint hypothesis that $\omega_{cc} = 0$ and $\omega_{xy} = 0$, it is natural to ask which equality is violated. Do covariance parameters differ across couples or over time, so that $\omega_{cc} > 0$? Or, are spouses' income volatilities correlated ($\omega_{xy} > 0$), so that husbands with large expected income shocks have wives with large expected income shocks? Section 3 showed how to use limited features of panel data to identify lower

Table 3: Sample Moments for Couples' Income Changes

Variables	Moment	Raw	Permanent	Transitory
$x_{it}y_{it}$	Mean	-0.0004	-0.0055	-0.0037
	Variance	0.0461	0.0692	0.0395
$x_{it}^2 y_{it}^2$	Mean	0.0461	0.0692	0.0396
	Variance	0.1955	0.2543	0.1279
x_{it}	Mean	0	0.0044	0.0014
x_{it}^2	Mean	0.3013	0.2970	0.2985
x_{it}^4	Mean	0.8560	0.8235	0.8473
y_{it}	Mean	0	0.0059	0.0002
y_{it}^2	Mean	0.1014	0.1983	0.0986
y_{it}^4	Mean	0.1024	0.2173	0.0954
$x_{it}y_{it}$ $x_{it-n}y_{it-n}$	Covariance	0.0009	0.0002	0.0013
x_{it}^2 y_{it-n}^2	Covariance	0.0017	-0.0008	0.0006
x_{it-n}^2 y_{it}^2	Covariance	0.0093	0.0190	0.0049
N	No. Obs.	20,762	15,478	19,430

This table presents sample moments that depend on x_{it} and y_{it}. For example, the sample moment for the mean of $x_{it}y_{it}$ is 1/NT $\sum_i \sum_t x_{it}y_{it}$ where N is the number of couples and T is the average number of observations per couple. For the "raw estimates" column, this is 1/NT $\sum_i \sum_t (z_{it}^x - z_{it-1}^x)(z_{it}^y - z_{it-1}^y)$. $x_{it} \equiv z_{it}^x - z_{it-1}^x$; $y_{it} \equiv z_{it}^y - z_{it-1}^y, n = 5$ if raw estimate; $y_{it} \equiv z_{it+2}^y - z_{it-3}^y, n = 6$ if permanent estimate; $y_{it} \equiv z_{it+1}^y - z_{it}^y, n = 5$ if transitory estimate. z_{it}^x is the Winsorized excess log income of the wife in household i in year t. z_{it}^y is the Winsorized excess log income of the husband in household i in year t. The construction of Winsorized excess log incomes is explained in the text. In brief, annual log labor incomes for husbands and wives are separately regressed on a host of covariates. The residuals from these regressions are Winsorized at the 5^{th} and 95^{th} percentiles.

Table 4: Testing for Correlated Heterogeneity:

Comparing the sample moment that identifies correlated heterogeneity to the distribution of this sample moment under null of no correlated heterogeneity

sample moment: $var(x_{it}y_{it})$

Raw	Permanent	Transitory
0.0461	0.0692	0.0395

Distribution of sample moment under null

Null 1	$\bar{c} = 0$; observations independent		
Mean	0.0305	0.0589	0.0294
(St.Dev.)	(0.0020)	(0.0034)	(0.0020)
$[z - stat]$	[7.6184]	[3.0552]	[4.9903]
Null 2	$\bar{c} \neq 0$ with normal shocks; observations independent		
Mean	0.0305	0.0589	0.0294
(St.Dev.)	(0.0031)	(0.0041)	(0.0026)
$[z - stat]$	[5.0732]	[2.5300]	[3.9413]
Null 3	"wife-swap bootstrap"; observations autocorrelated but independent of spouse's		
Mean	0.0296	0.0573	0.0286
(St.Dev.)	(0.0022)	(0.0036)	(0.0021)
$[z - stat]$	[7.68]	[3.28]	[5.010]

$x_{it} \equiv z^x_{i,t} - z^x_{i,t-1}$; $y_{it} \equiv z^y_{i,t} - z^y_{i,t-1}$ if raw estimate; $y_{it} \equiv z^y_{i,t+2} - z^y_{i,t-3}$ if permanent estimate; $y_{it} \equiv z^y_{i,t+1} - z^y_{i,t}$ if transitory estimate. The null hypothesis is that there is no correlated heterogeneity, formally the joint hypothesis that $\omega_{cc} = \omega_{xy} = 0$. $\omega_{cc} = 0$ means that the covariances between spouses' income changes (expected product of husbands' and wives' changes in excess log income) are the same for all i and t. $\omega_{xy} = 0$ means that differences over i and t in husbands' and wives' volatilities (expected squared changes in excess log income) are uncorrelated. the covariance between x_{it} and y_{it} is the same for all i and t. The three nulls listed identify additional assumptions about the joint distribution of x_{it} and y_{it}. Under Null 1, x_{it} and y_{it} are uncorrelated on average and x_{it} and y_{it} are independent of x_{js} and y_{js} if $i \neq j$ or $t \neq s$. Under Null 2, x_{it} and y_{it} have the average covariance estimated from the data and shown in Table 3, the correlated component of shocks has a normal distribution, and x_{it} and y_{it} are independent of x_{js} and y_{js} if $i \neq j$ or $t \neq s$. Under Null 3, the "wife-swap bootstrap" explained in the text is used to build the reference distribution under the null. This implicitly assumes that x_{it} is independent of y_{js} for all j and s including $j = i$ and $s = t$, that x_{it} and y_{it} may be correlated with x_{is} and y_{is}, respectively, and have correlations taken from the data, but that x_{it} and y_{it} are uncorrelated with x_{js} and y_{js}, respectively, if $i \neq j$.

bounds on ω_{cc} and ω_{xy}. As shown in Section 3, these bounds exploit the intuition that covariances must be less than variances in the sense that $|cov(a_t, a_{t-n})| < var(a_t)$. A lower bound on ω_{xy} implies an upper bound on ω_{cc}, and the two lower bounds imply an upper bound on the tail-fatness of correlated shocks.

So that these bounds usefully identify heterogeneity in parameters and not autocorrelated shocks, I calculate covariances based on income changes that are sufficiently far apart that their shocks are conditionally uncorrelated. Abowd and Card (1989) show that one-year income changes that are two or more apart are uncorrelated. Given the large number of years of data spanned by the transitory and permanent covariance estimators, I use relatively large distances between estimates. I set $n = 5$ for the raw and transitory covariance, so that I examine the sample covariance of the product of couples income changes with the product of their income changes five years ago. For the permanent covariance (the estimator of which spans 5 years of data), I set $n = 6$.

The formulas used to calculate these bounds are desctibed in Section 3; the sample moments that are plugged into these formulas are shown in Table 3. These bounds are shown in Table 5. These bounds allow us decompose sample variation in the proxy for a couples' covariance parameter, $x_i y_i$, as (repeated from equation 3 with $\bar{c} = 0$):

$$var(x_i y_i) = \bar{\sigma}_x^2 \bar{\sigma}_y^2 + \omega_{xy} + \omega_{cc}(\kappa_c - 1) \tag{9}$$

Note that two-thirds (in the case of the raw estimates) or more (in the case of the permanent and transitory estimates) of the variation in the covariance parameter proxy, $x_i y_i$, can be explained by realized shocks. This is shown in the second row of data, which shows how much variation in $x_i y_i$ would be expected even if x_i and y_i evolved independently.

However, there is also evidence of parameter heterogeneity. For the raw covariance, $0.001 < \omega_{cc} < 0.005$. The upper bound will be brought down if correlated shocks have tails that are fatter than a normal distributions'. This implies a reliabililty ratio of between 2 and 11 percent. In other words, of the variation in $x_{it} y_{it}$ (where x_{it} and y_{it} are contemporaneous one-year changes in excess log income for the wife and her husband), between 2 and 12 percent is differences in covariance parameters across couples or over time. Given equation 9, variation in couples' covariance parameters, ω_{cc}, will explain between $2(\kappa_c - 1)$ and $12(\kappa_c - 1)$ percent of the sample variation in the proxy for couples' covariance paramter, $x_{it} y_{it}$ (between 4 and 24 percent when shocks are normal so that $\kappa_c = 3$). The remainder reflects differences across observations in realized shocks. While the lower bound (2 percent) can be increased by reducing n , the time gap between observations used in the covariance, below 5, this raises the risk that bounds are contaminated by correlated shocks. For the transitory covariance, the bounds imply a reliability ratio of between 3 and 7 percent.

In the case of the permanent covariance, correlated heterogeneity can be explained by couples having correlated volatility, ω_{xy}, and not by different couples having different covariances of permanent income shocks, ω_{cc}. This is apparent because there is little covariance between the covariance measure $x_{it} y_{it}$ and its lag (so that the

Table 5: Decomposing Correlated Heterogeneity

Percent of variation in "estimated covariance", $x_i y_i$, attributed to covariance heterogeneity, correlated variances, or realized shocks without either

Decomposition: $var\left(x_i y_i\right) = \bar{\sigma}_x^2 \bar{\sigma}_y^2 + \omega_{xy} + \omega_{cc}\left(\kappa_c - 1\right)$

	Raw Estimates	Permanent Estimates	Transitory Estimates
$var(x_{it} y_{it})$	0.0461	0.0692	0.0395
variation when x and y are independent, $\overline{\sigma}_x^2 \overline{\sigma}_y^2$	0.0305	0.0589	0.0294
(% of variation)	(66.2%)	(85.1%)	(74.4%)
correlated variances, $\underline{\omega_{xy}}$	> 0.0055	> 0.0091	> 0.0028
(% of variation)	($> 11.9\%$)	($> 13.1\%$)	($> 7.0\%$)
covariance heterogeneity, $\underline{\omega_{cc}}$	> 0.0009	> 0.0002	> 0.0013
(% of variation)	($> 2.0\%$)	($> 0.3\%$)	($> 3.2\%$)
covariance heterogeneity, $\overline{\omega_{cc}}$	< 0.0051	< 0.0006	< 0.0037
(% of variation, $\kappa_c = 3$)	($< 11.0\%$)	($< 0.8\%$)	($< 9.3\%$)
kurtosis, $\overline{\kappa_c}$	< 12.2	< 6.50	< 6.8

For raw estimates, x_{it} and y_{it} are contemporaneous one-year changes in excess log incomes for wives and their husbands; for permanent estimates, x_{it} are one-year income changes for wives while y_{it} are the five-year changes for their husbands that span their wives' short-term changes; for transitory estimates, x_{it} are one-year income changes for wives while y_{it} are the one-year changes for their husbands lagged one year. $var(x_{it} y_{it})$ indicates the sample variance of $x_{it} y_{it}$ in each case. $\overline{\sigma}_x^2 \overline{\sigma}_y^2$ is the product of the sample variances of x_{it} and y_{it}, which is the expected value of $var(x_{it} y_{it})$ when x_{it} and y_{it} are independent. $\underline{\omega_{xy}}$ is the lower bound on correlated variances, the covariance between husbands' and wives' variances, $cov(\sigma_{xit}^2, \sigma_{yit}^2)$. $\underline{\omega_{cc}}$ and $\overline{\omega_{cc}}$ are lower and upper bounds on covariance heterogeneity, the variance of the covariance parameter governing couples' shocks. $\overline{\kappa_c}$ is an upper bound on the kurtosis of couples' correlated income shocks. See text for construction of these bounds.

lower bound on ω_{cc} is close to zero), while there is substantial covariance between one spouse's volatility and the lagged volatility of his or her partner (so that the lower bound on ω_{xy} is substantial and therefore the upper bound on ω_{cc} is small as well).

There is also evidence of correlated heterogeneity, so that couples in which the husband's income has a high variance parameter also tend to have a wife whose variance paramter is also high. This explains at least 12 percent of the variation in $x_i y_i$ in the case of raw estimates (and more than 13 percent and more than 7 percent for the permanent and transitory estimates, respectively). This is identified from the high correlation between one spouse's squared changes in income and the lagged squared changes in their partner's income. Note that these lower bounds on ω_{xy} are what generate an upper bound on ω_{cc}. Put simply, the more variation in the data can be explained by correlated variances, ω_{xy}, the less is left over to be explained by covariance heterogeneity, ω_{cc}. However, the trade-off between these depends on the kurtosis of correlated shocks, κ_c, so kurtosis must be guessed or assumed to obtain an upper bound on ω_{cc}.

The lower bounds on ω_{cc} and ω_{xy} imply an upper bound on the kurtosis of correlated shocks, κ_c. If large variation in $x_{it}y_{it}$ is attributable to only a tiny amount of heterogeneity but extremely fat-tailed correlated shocks, then $x_{it}y_{it}$ should not be with its lags so long as they are far enough apart. If this autocorrelation is substantial, tails of correlated shocks cannot be too fat. Using this method implies that kurtosis must be less than 13 for correlated raw shocks and less than 5 for correlated transitory ones.

Upper and lower bounds on ω_{cc} imply an upper and lower bound on the reliability ratio. $x_{it}y_{it}$ can be included as a right-hand-side variable in a regression as a proxy for covariance. The attenuation bias that is introduced by including this noisy measure can be corrected explicitly with the reliability ratio. This can be corrected if $x_{it}y_{it}$ is multiplied by $\omega_{cc}/var\,(x_i y_i)$. While useful when using a raw or transitory estimate $x_{it}y_{it}$ as a proxy for the couples' covariance parameter, c_i, this will not be fruitful in the case for permanent estimates. Almost none of the variation in permanent covariance measures ($x_{it}y_{it}$) reflects heterogeneity in the permanent covariance (ω_{cc}).

Section 3.2 outlined how heterogeneity could be identified if there are covariates than can explain some of the variation in $x_{it}y_{it}$ or the correlated component of x_{it}^2 and y_{it}^2. In the case of couples' joint income processes, the most powerful such variable is the number of years the couple has been married. Shore (2006) shows that couples' shocks are most negatively correlated early in marriage and become increasingly positively correlated throughout the life of a marriage. While this pattern is strong statistically (the hypothesis that there is no time variation can be rejected at the 0.1 percent level) and economically (with correlations increasing from as low as -30 percent to as high as 10 percent), time patterns explains less than 1 percent of the variation in $x_{it}y_{it}$ or the correlated variation in x_{it}^2 and y_{it}^2. In this instance, covariates are of only limited help in placing a lower bound on ω_{cc}.

5 Conclusion

While identifying latent heterogeneity presents substantial challenges to econometricians, it is possible to test for latent correlated heterogeneity as long as covariances are close to zero. Unless one is willing to assume the distribution of shocks or error (e.g., kurtosis), it is generally impossible to differentiate heterogeneity in a parameter from error in the estimation of that parameter. This paper shows that when covariances are zero on average, a rejection of independence using second moments can be interpreted as a rejection of the hypothesis that there is no correlated heterogeneity. Unlike the case of heterogeneity in a single variable, the analogous term for correlated heterogeneity that is affected by the shape of the distribution drops out.

Correlated heterogeneity can be decomposed into heterogeneity in covariances or correlated heterogeneity in variances. This paper shows how to use additional information (e.g., time-series data) to place lower bounds on the relative importance of these two. These approaches mirror the approaches to use additional information to place a lower bound on heterogeneity in a single variable. But in the case of correlated heterogeneity, a lower bound placed on covariance heterogeneity places an upper bound on correlated variance heterogeneity, and *vise versa*. While these upper bounds require as assumption about tail-fatness, results are much less sensitive to tail-fatness mis-specification than in the analogous case for heterogeneity in a single variables.

Identifying correlated heterogeneity is of substantial value. First, estimates of the covariance on average may not be sufficient for welfare calculations when a representative-agent model is inappropriate. For example, if couples select into various covariances based on their risk preferences (with risk-averse couples choosing lower covariances), knowing the average level of risk aversion and the average covariance will understate the welfare benefits of risk-sharing by couples. Second, there are a variety of settings in which covariances are useful as right-hand-side variables. For example, Hess (2004) uses couples' covariances to predict divorce as a test of competing theories of marriage; Shore (2007) uses them to predict saving and consumption as a test of precautionary saving. Since instruments for couples' covariances are weak (and of dubious exogeneity), it is more fruitful to exploit the full range of variation in covariances in the data. To correct for the attenuation bias caused by including noisy measures of covariance as right-hand-side variables, we need the fraction of variation in parameter estimates that stems from variation in parameters (as opposed to estimation error).

References

Abowd, J., and D. Card (1989): "On the Covariance Structure of Earnings and Hours Changes," *Econometrica*, 57(2), 411–445.

Alvarez, J., M. Browning, and M. Ejrnaes (2001): "Modelling income processes with lots of heterogeneity," University of Copenhagen Working Paper.

Browning, M., and A. Lusardi (1996): "Household Saving: Micro Theory and Micro Facts," *Journal of Economic Literature*, 34, 1797–1855.

Carroll, C. D., and A. A. Samwick (1997): "The Nature of Precautionary Wealth," *Journal of Monetary Economics*, 40(1), 41–71.

Hess, G. D. (2004): "Marriage and Consumption Insurance: What's Love Got to Do with It?," *Journal of Political Economy*, 112(2), 290–318.

Jensen, S. T., and S. H. Shore (2007): "The Evolution of the Distribution of Income Volitility," Working Paper, Johns Hopkins University.

Meghir, C., and L. Pistaferri (2004): "Income Variance Dynamics and Heterogeneity," *Econometrica*, 72(1), 1–32.

Schulhofer-Wohl, S. (2006): "Heterogeneity, Risk Sharing and the Welfare Costs of Risk," Working Paper.

Shore, S. H. (2006): "The Co-Movement of Couples' Incomes," University of Pennsylvania Working Paper.

——— (2007): "Couples' Co-Movement and Precautionary Saving," Johns Hopkins University.