

## Tilburg University

### Time, money, peers, and parents

Kooreman, P.

*Published in:*  
Journal of Population Economics

*Publication date:*  
2007

[Link to publication in Tilburg University Research Portal](#)

*Citation for published version (APA):*  
Kooreman, P. (2007). Time, money, peers, and parents: Some data and theories on teenage behavior. *Journal of Population Economics*, 20, 9-33.

#### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

# Time, Money, Peers, and Parents; Some Data and Theories on Teenage Behavior

Peter Kooreman \*

University of Groningen

Version December 2004

## Abstract

In the first part of the paper I analyze a data set on teenage behavior. The data is a sample of high school students in the Netherlands, and contains information on teenage time use, income, expenditures, and subjective measures of well-being and self-esteem. As all students in a sampled class are interviewed in principle, the data set has rich information on the behavior of potentially important peers of each respondent. I estimate models to assess (bounds on) the magnitude of endogenous social interactions. For some types of behavior (*e.g.* truancy, smoking, pocket money, alcohol expenditures), endogenous social interactions within school classes are strong; for other behaviors they are moderate or unimportant. Within-gender interactions are generally stronger than interactions between boys and girls, with some intriguing exceptions.

In the second part of the paper I discuss a number of theories that might help to understand the empirical patterns. Key concepts in the discussion are interdependent preferences, endogenous social norms, identity, and intergenerational interactions.

*Keywords:* teenage behavior, peer effects, time use, expenditures

---

\*Department of Economics, University of Groningen, P.O. Box 800, 9700 AV Groningen, The Netherlands; fax: +31 50 363 7337; email: p.kooreman@eco.rug.nl. Revised version of the Presidential Address to the Seventeenth Annual Conference of the *European Society for Population Economics (ESPE)*, New York City, June 15, 2003. I thank participants at the ESPE 2003 Conference, two anonymous referees and the editor of this journal, as well as Daniel Hamermesh, Joyce Jacobsen, and Adriaan Soetevent for helpful comments.

*JEL classification:* D12, D13

# 1 Introduction

Economic aspects of teenage behavior can be studied from several perspectives and with different motivations.<sup>1</sup> First, teenagers can be interesting from a *marketing* perspective. Total expenditures by teenagers amount to several percentages of GDP (*cf.* Warnaar and Van Praag, 1997) and in some markets (*e.g.* designer clothing, cell phones, video games) teenagers are important customers.

A second perspective (perhaps more appealing to population economists), is *public policy*. Certain aspects of teenage behavior are a source of concern to policy makers, teachers and parents: using drugs, dropping out of school, teenage pregnancy, smoking cigarettes, and drinking alcohol are examples. These behaviors bring health risks (such as alcohol related traffic accidents), large social costs, and may have long-lasting effects on individuals. Effective policies to influence these behaviors require a thorough understanding of what teenagers prefer, what resources they have available, and how they make their decisions.

Thirdly, teenage behavior is a natural area for empirical analyses of social interactions, and therefore interesting from an intrinsic *research* perspective. While teenage behavior is obviously also influenced by persons outside the class, classmates are likely to play a potentially dominant role in shaping teenagers' preferences and behavior. In this paper I will use a school class based data set in which all students within a sampled class are interviewed in principle. As a result, the data set has unusually rich information on members of a sampled individual's reference group. This is in sharp contrast with the situation when analyzing household surveys, which typically lack explicit questions on the behavior of households in a household's reference group.

The paper proceeds as follows. Section 2 describes the data of the Dutch National School Youth Survey. The survey contains information on teenage time use, income, expenditures, and subjective measures of well-being and self-esteem. Section 3 is an exploratory analysis of the data and estimates simple reduced form models to measure the effect of explanatory variables as well as the strength of the within-class correlation of student behavior. Section 4 reconsiders identification issues and estimates a model that explicitly allows for endogenous social interactions. Due to the cross-sectional nature of the data I cannot fully control for all possible biases in the measurement of endogenous social interaction effects. Although I will argue that these biases are likely to be small, the estimated effects should be interpreted as upper bounds on the true endogenous social interaction effects.

I find that for some aspects of teenage behavior (*e.g.* truancy, smoking, pocket money, alcohol expenditures) social interaction effects are strong, while they are moderate or unimportant for other aspects of behavior. I also estimate a version of the model that distinguishes between within-gender and between-gender interactions. Within-gender interactions are generally stronger than interactions between boys and girls, with some intriguing exceptions. Having jobs, going-out, smoking, drinking, and truancy show similar empirical patterns suggesting that the concurrence of these behaviors define a life style or identity. Sections 2, 3, and 4 fit into an applied econometrics literature that focuses on the measurement of social interactions (see Duflo and Saez (2003), Durlauf and Moffitt (2003) and Kawaguchi (2004) for some recent contributions).

In section 5, I review theories that may help to put the empirical results into perspective. Key concepts are interdependent preferences, social norms, identity, and intergenerational interactions. This section is related to a public

economic theory literature, and also borrows from social psychology. A sketch of possible routes for future research in section 6 concludes the paper.<sup>2</sup>

## 2 The Dutch National School Youth Survey

The empirical analysis is based on the Dutch National School Youth Survey (NSYS). The first NSYS survey took place in 1984 (ESPE's first president Bernard van Praag being one of the initiators). The survey was held for the second time in 1990, and has been repeated bi-annually since then. The survey is a joint effort of the Social and Cultural Planning Office of The Netherlands (SCP) and the Netherlands Institute for Family Finance Information (NIBUD). Each survey is based on a random sample of some 500 high school classes with approximately 12,000 students. A school that participates is compensated by means of a report summarizing the survey results for that school. The series of surveys is not a panel, although some schools have participated more than once.

In principle all students in a sampled class participate in the survey. Yet, some of them are excluded from the data, for example because a student was absent on the day when the questionnaires were filled out. The survey contains a wealth of information on economic, social, and psychological aspects of teenage life. More specifically, I will look at how teenagers spend their time, how they get money and how they spend it, and how they assess their self-esteem and well-being, mainly using the 1992 data.<sup>3</sup> There is limited information on parents (education and working hours) and on siblings.

A US data set which is comparable to the present one is the National Education and Longitudinal Study (NELS), see *e.g.* Gaviria and Raphael (2001). Both the Dutch NSYS and the NELS focus on non-cognitive outcomes within schools. The NELS is a biannual survey, first held in 1988, and

samples students within roughly 1000 schools. An important difference with the Dutch NSYS is that the NELS surveys only a relatively small group of students within each school. For example, in the 1990 sample used by Gaviria and Raphael, the mean sample size per school was 13.3 students. While the NELS contains information on school averages, these are not available per class, grade, or gender. This limits the possibilities for an analysis of interactions within schools (for example, it is impossible to allow for a school specific fixed effect) and it precludes any analysis of social interactions within classes. Two other US data sets on teenagers with peer group information are the Teenage Attitudes and Practices (TAPS) and the National Longitudinal Survey of Youths (NLSY). However, the TAPS only contains subjective information on a respondent's four best same-sex friends, whereas the NLSY only has subjective peer information based on questions of the type "*What percentage of kids in your grade...?*".

Table 1 presents a number of sample statistics for the 1992 data. The three by far most important time use categories for teenagers are sleep (59 hours per week, 8.4 hours per day), school (38 hours per week), and "screens" (a summary term for time spent on watching TV and video and using a computer; 21 hours per week, 3 hours per day). Other categories are very much smaller. The numbers refer to time spent during a "normal school week" (thus "jobs" excludes time spent on vacation jobs). The expenditure figures refer to the question "*How much do you spend on ... from your own money*" (thus, expenditures on teenage clothing directly by parents, for example, are excluded). The measure of self-esteem is based on the following six questions: 1) "*I sometimes think I will never be good at anything*", 2) "*People like me have little chance to become successful in life*", 3) "*I think others find me ugly*", 4) "*I feel very insecure when I have to make a decision*".

by myself ", 5) "I'm often afraid to fail ", and 6) "I often feel unsure when I'm in company with others ". The measure is defined as six minus the number of affirmative answers to these questions. Thus the measure ranges from 0 (affirmative answers only) to 6 (no affirmative answers). The measure of well-being is based on the question: *Considering everything, how do you feel you are doing now?* The possible answers (*very good, good, fair, poor, bad, very bad*) have been linked to the numbers 6 to 1, respectively.

One difficulty with the data is caused by the fact that a large proportion of teenagers – about 30 percent – do not know their parents' education level. Separating these percentages for boys and girls shows a remarkable gender pattern of interactions between parents and children. With respect to father's education sons appear to be better informed than daughters (72 versus 66 percent). For mother's education I find the opposite: daughters are better informed than sons (72 versus 68 percent). These results are in line with those on gender patterns in parent-child interactions reported by Thomas (1990). Given this data feature, one option is to exclude observations with missing information on parental education levels. Another one is to include these observations and perform an analysis without parental education as explanatory variables. The empirical results reported below are based on the first option.

### **3 Reduced form regressions**

As a first step in the data analysis I have run a regression for each of the eighteen dependent variables: ten time use categories (sleep, eating/personal care, school, jobs, household chores, going-out, "screens", reading, music, and sports), two income variables (income from parents, income from jobs), three expenditure categories (clothing and shoes, cosmetics, and alcohol), savings,



and two subjective measures (self-esteem and well-being).

The list of explanatory variables is largely determined by data availability: a constant, gender, age, non-dutch, single-parent family, family size, urbanization, the student's school level, father's and mother's education, father's and mother's weekly working time, and religion. Unfortunately, there is no direct measure of family income in the data. As a consequence, coefficients on father's and mother's education and hours of work may partly reflect income effects.

Let  $x_{ik}$  be a row vector of observable exogenous variables for individual  $i$  in class  $k$ ,  $i = 1, \dots, N_k$ , where  $N_k$  is the number of students in class  $k$ ;  $y_{ik}$  is the dependent variable. Let  $\epsilon_{ik}$  be an error term representing all unobserved explanatory variables;  $\epsilon_{ik}$  is assumed to be independent of all exogenous variables, and  $E\epsilon_{ik} = 0$ . The standard regression model

$$y_{ik} = x_{ik}\beta + \epsilon_{ik} \tag{1}$$

is extended slightly by allowing the error terms of students within a class to be correlated:

$$\Omega \equiv E\epsilon_k\epsilon_k' = \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \dots & 1 \end{bmatrix},$$

where  $\epsilon_k = (\epsilon_1, \dots, \epsilon_{N_k})'$ . In addition, I assume  $\text{cov}(\epsilon_k, \epsilon_\ell) = 0, \dots, k \neq \ell$ , i.e. error terms of two students who are from different classes are assumed to be 0.<sup>4</sup>

A significant correlation coefficient indicates the presence of some kind of interdependence in the behavior of classmates. There may be unmeasured explanatory variables that affect each student within a class, for example the

behavior of teachers, school policies, or similar family and cognitive backgrounds that have selected the students into the same class. There may also be a correlation if there are endogenous social effects (that have not yet been modeled explicitly).

### **3.1 Results: explanatory variables**

Before looking at the estimates of the within-class correlation coefficients in detail, I discuss the effects of the explanatory variables. I choose to do so by explanatory variable (tables 2, 3, 4, and 5).

#### *Gender*

Girls spend more time on school, household chores, sleep, and personal care and eating than boys do. The differences, however, are not very large: at most one hour per week per category (but slightly more than one hour per week for time spent on school). The time use categories on which girls (necessarily) spend less time than boys are jobs, sports, and time spent on “screens” (TV, video, and computer). The difference for the latter category is large and highly significant: boys spend two hours per week more in front of screens than do girls.

Parents do not discriminate between boys and girls in terms of the amount of money they give to their children: the girl dummy is insignificant in the income-from-parents regression. But it is not in the income-from-jobs regression. Girls earn about 25 percent less than boys from jobs, slightly more than the difference in time spent on jobs. Yet, girls spend 50 percent more on clothing than boys, a gender effect also found for single adults in a different study (Kooreman, 2000). Expenditures on alcohol also show a very strong gender effect, but now with the opposite sign: boys spend twice as

much on alcohol than girls. Alcohol expenditures are substantially lower for non-Dutch teenagers. Girls save about 30 percent less than boys.

Gender has a large effect on self-esteem. On the 0 to 6 scale, girls have a 0.5 lower score of self-esteem, and the effect is statistically very significant. The effect of gender on well-being is not significant. (Self-esteem and well-being are obviously related, but well-being is a broader concept that involves more than self-esteem.)

### *Single-parent families*

Teenagers in single parent households spend (much) more time on household chores (almost two hours per week more), but also on personal care and going out. This seems to come at the expense of sleep. Teenagers from a single-parent family get about 50 percent more money from the parent than teenagers in two-parent families. Perhaps this indicates that a teenager in a single parent family is required or supposed to behave more independently than peers from otherwise identical families. Alternatively, it might be a compensation for the time spent on household chores. This interpretation would imply an implicit wage rate of 6 guilders for an extra hour of household chores.<sup>5</sup> Girls in single-parent families spend more on clothing and cosmetics than girls in otherwise identical two-parent families. Being in a single parent family also has a significant negative effect on well-being, but there is no effect on self-esteem. Living with only a single parent also increases the teenager's probability to own a cell phone.<sup>6</sup>

### *School level*

The school level has a very large impact on time use. Compared to students in a vocational program, students in VWO (giving access to college and university) spend four hours per week more on school. They spend less

time on jobs, household chores, going out, screens, and sleep. School level does not have an effect on the amount of money teenagers receive from their parents, but income from jobs strongly decreases with school level (consistent with the result found for time spent on jobs). This effect on income is reflected in expenditure patterns, in particular alcohol expenditure, savings, and moped ownership. An intriguing result is that school level has a positive effect on self-esteem, and a negative effect on well-being. Apparently, a higher school level has a benefit (self-esteem), but the price is high (time spent on school).

*Father's and mother's hours of work and education level*

Father's and mother's hours of work have a positive effect on the time the teenager spends on going out, on jobs, and on alcohol expenditures. This may represent an income effect, or it might be related to lower parental attention. For a number of other behaviors I do find a significant effect of mother's working hours and insignificant effects for father's working hours. Examples are expenditures on clothing, cosmetics, and savings, and income from parents. These differences again may point at a labeling effect: The marginal propensity to spend money on the children out of mother's income is (much) larger than the marginal propensity to spend money on the children out of father's income (*cf.* Kooreman, 2000). Similar asymmetric effects of father's and mother's working time are found for truancy, smoking, cell phone ownership, and time spent on household chores. The insignificance of the effect of father's working time may be related to its smaller variance in the data set.

Variables that have a strong positive effect on self-esteem are the student's school level and the father's education level. The effect of the mother's

education level on self-esteem is not significant, and interaction coefficients (not reported) indicate that the effects of parental education level are the same for boys and girls.

### *Family size*

Students from large families spend more time on jobs and household chores, and less on screens and sports. Children from smaller families have higher self-esteem.

### *Urbanization*

Urbanization (ranging from 1 (large city) to 5 (rural area)) has a significant effect on time spent on TV, video, and computer: a teenager in a large city spends almost two hours per week more on screens than a teenager in a rural area. Teenagers in large cities spend more on clothing and shoes than those in rural areas.

### *Religion*

Religion does not have much of an effect on time use. An exception is the negative effect of being Protestant on time spent on screens and the positive effect on reading.

### *Age*

Both income sources strongly increase with age, especially income from jobs. There is a significant decrease in well-being when the teenager grows older, but self-esteem – on the other hand – does not relate to age.

## 3.2 Results: within-class correlations

We now turn to the estimated within-class correlation coefficients for the various time use categories. The largest correlation coefficient – the one for time spent on school – is 0.12, which means that 12 percent of the unexplained variance in the regressions can be attributed to factors common to all students within the class. Of course, it is plausible to find a positive correlation for an activity closely related to school. Other categories where I find significant within-class correlations are jobs, going out, “screens”, and sports. The significant positive correlation for sleep might be due to the time adding-up constraint. A more interesting question is whether there is also a correlation for activities unrelated or remotely related to school. A correlation in that case could be an indication for selection effects into classes related to a similarity in socioeconomic backgrounds of families. However, for time use categories far away from school I find the correlations to be essentially zero. Examples are household chores, personal care, music, and reading.

For both income sources there is a significant within-class correlation coefficient, but the coefficient for income from parents is twice as large as the coefficient for income from jobs (0.059 versus 0.029).

Looking at the within-class correlations for expenditures, the largest correlation is found for alcohol expenditures, 0.056, and the second largest for clothing expenditures, 0.019. Both these correlation coefficients are significant. The correlation coefficient for savings is insignificant.

The intra-class correlation coefficient is positive and significant for self-esteem and insignificant for well-being. These results suggest that self-esteem is primarily tied to the school and the class, whereas well-being is also related to the family and the home environment.

## 4 Endogenous social interactions

In order to get a better picture of the possible mechanisms behind these correlations, I estimate a model which explicitly allows for endogenous social interactions. Section 4.1 presents the model and reconsiders some identification issues, followed by the presentation of the empirical results in section 4.2. Section 4.3 extends earlier work by making a distinction between interactions within and across genders.

### 4.1 Some identification issues reconsidered

Consider the model

$$y_{ik} = x_{ik}\beta + \gamma\bar{y}_{-ik} + \bar{x}_{-ik}\delta + \epsilon_{ik} \quad (2)$$

where

$$\bar{y}_{-ik} = \frac{1}{N_k - 1} \sum_{\substack{j=1 \\ j \neq i}}^{N_k} y_{jk} \quad \bar{x}_{-ik} = \frac{1}{N_k - 1} \sum_{\substack{j=1 \\ j \neq i}}^{N_k} x_{jk}$$

A positive  $\gamma$  reflects an inclination to conform to the behavior of others, a negative  $\gamma$  an inclination to deviate from the behavior of others.

It turns out convenient to rewrite (1) in matrix notation as

$$y_k = X_k\beta + \Gamma y_k + \tilde{X}_k\delta + \epsilon_k$$

with

$$\Gamma = \begin{bmatrix} 0 & \frac{\gamma}{N_k-1} & \cdots & \frac{\gamma}{N_k-1} \\ \frac{\gamma}{N_k-1} & 0 & \cdots & \frac{\gamma}{N_k-1} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\gamma}{N_k-1} & \frac{\gamma}{N_k-1} & \cdots & 0 \end{bmatrix}.$$

Let  $A \equiv (I - \Gamma)^{-1}$ . The reduced form of (2),

$$y_k = AX_k\beta + A\tilde{X}_k\delta + A\epsilon_k, \quad (3)$$

will be the basis for estimation. Note from (3) that if  $\gamma \neq 0$ , the reduced form error terms within a class (i.e. elements of  $v_k \equiv A\epsilon_k$ ) will be correlated even if  $\rho = 0$ .

The estimation algorithm explicitly evaluates the reduced form (the social equilibrium). The parameters are then estimated by maximum likelihood, assuming that each  $\epsilon_k$  follows a multivariate normal distribution with the same correlation structure as in section 3.

A number of papers, notably those by Manski (1993 and 2000) describe potential caveats in the identification of social interactions. Given the nature of the present data I will not be able to fully account for all of these problems. In order to provide a proper perspective for the interpretation of the empirical results to be presented, I will discuss identification problems in relation to the present data set: *i)* the definition of the reference group, *ii)* non-random selection into reference groups, *iii)* simultaneity of mutual endogenous interaction effects, and *iv)* the discrete nature of some of the endogenous variables.

*The definition of the reference group*

One of the methodological problems in analyses of social interactions is how to define a reference (social) group: the group of individuals who affect the behavior of a given individual. A number of authors have defined the reference group of an individual as the group of all persons in the population within the same age group and with the same education level, and used the sample analogues as an approximation; see *e.g.* Kapteyn *et. al* (1997) and Aronsson *et al.* (1999). This is a crude definition, largely motivated by data limitations. A more attractive alternative is to use subjective in-



formation on an individual's reference group, as in Woittiez and Kapteyn (1998). However, the information on the reference group of a sampled individual is often limited as these reference group members are not themselves included in the sample. The data in the current analysis can be viewed as a *reference group based sample* as all students within a sampled class are interviewed in principle. While teenage behavior is obviously also influenced by persons outside the class, classmates play a potentially dominant role in shaping teenagers' preferences and behavior. On a weekday, the average student spends about six hours in his or her school class. The total time spent on school related activities (including homework and commuting) is about eight hours per weekday, more than fifty percent of the daily waking time. Teenagers within the same school or class therefore form social groups that are more clearly defined and delineated than in many other situations in which social interactions are likely to play a role.

#### *Non-random selection into reference groups*

In The Netherlands high schools are state-funded and there are no restrictions in terms of catchment areas. In practice, school choice is based on geographic proximity, (religious) denomination, type of education (level and pedagogical principles), and (perceived) school quality. Conditional on school choice, selection into classes is primarily based on cognitive abilities. In the regressions reported below, the level of education is one of the explanatory variables. In combination with allowing for within-class correlation of error terms, this controls for non-random selection into classes to some extent. I also report some results on models with school specific fixed effects, which controls for non-random selection into school to some extent.

#### *Simultaneity of mutual endogenous social interaction effects*

Equation (2) can be interpreted as a reaction function. Due to its linearity there is a unique (Nash) equilibrium (provided that  $I - \Gamma$  is non-singular). The identification problem is now is whether one can uniquely solve the structural parameters  $\beta$ ,  $\gamma$ , and  $\delta$  from the reduced form parameters. It is well-known that this is generally impossible without imposing a priori restrictions on the parameters. Without such restrictions, one can identify whether social interactions are present, but it is impossible to distinguish between endogenous social interactions and social interactions of other types.

In practical applications there often are credible parameter restrictions. For example, it is difficult to see why the income of the parents of student  $j$  would have a *direct* effect on the pocket money student  $i$  receives from her parents. If we would find, in a reduced form analysis, that there is a significant effect of the income of the parents of student  $j$  on the amount of pocket money student  $i$  receives from her parents, then the only plausible explanation seems to be that this is an indirect effect that runs via the pocket money of student  $j$ . Note also that for identification it is sufficient that there is only a single exogenous variable of student  $j$  that does not have a direct effect on the dependent variable of student  $i$ .

Gaviria and Raphael (2001) argue that students are less exposed to the family background of their school peers than they are exposed to the family background of peers residing in the same neighborhood. They conjecture that in an analysis of interactions through schools contextual effects (measured by the parameter  $\delta$  in equation (2)) are less important than in an analysis of interactions through neighborhoods. In their empirical analysis they assume that contextual effects are absent ( $\delta = 0$ ). Kawaguchi (2003) invokes subjective information about the perception of peer behaviors to achieve full identification.<sup>7</sup> He finds that the absence of contextual effects ( $\delta = 0$ ) cannot

be rejected. The empirical results presented below are based on the assumption that there are no contextual effects. The estimates on the endogenous social interaction effects should therefore be interpreted as upper bounds on the true effects.<sup>8</sup>

#### *Discrete nature of endogenous variables*

Several of the endogenous variables I consider are discrete (or mixed discrete-continuous), *e.g.* cell phone ownership, smoking, and drinking alcohol. In these cases there may exist multiple (Nash) equilibria. The estimation procedure used in those cases is described in Soetevent and Kooreman (2004).

## **4.2 Results**

Regarding time use strong endogenous interactions are found for time spent on school and for jobs (table 6). Both coefficients are close to 0.25, implying a multiplier of 1.33 (*i.e.* if some exogenous force initially increases the time each student spends on school by  $x$  percent, then the increase in time spent on school in equilibrium will be 1.33 times  $x$  on average, due to social interactions; similarly for jobs). The endogenous interaction effects are also significant for time spent on “screens”, sports, and going-out. The coefficients are about 0.15, implying a multiplier of 1.18. The two largest of all estimated endogenous interaction coefficients are those for income from parents (0.341,  $t$ -value 5.8) and alcohol expenditures (0.309,  $t$ -value 7.0). The implied multipliers are 1.52 and 1.45, respectively. Significant but smaller effects are also found for income from jobs and for expenditures on clothing. For savings the endogenous interaction coefficient is not significant.

The only negative estimated endogenous interaction coefficient is found

for self-esteem, suggesting that self-esteem of a student goes at the expense of the self-esteem of others in the class. The coefficient is not significant, however. For well-being the coefficient is 0.171 ( $t$ -value 2.7), implying a multiplier of 1.21.

### 4.3 Within-gender versus cross-gender interactions

In the most general specification, interactions within a class are represented by an arbitrary  $N_k \times N_k$  interaction matrix  $\Gamma$ , where element  $(i, j)$  measures the effect of student  $j$  on student  $i$  (diagonal elements are 0). The results presented above were based on the assumption that all non-diagonal elements were equal and dependent on only a single parameter,  $\gamma$ . Obviously, there are possibilities for relaxing this symmetry assumption. A first one that comes to mind is a refinement based on the boy-girl distinction. Then the elements of the interaction matrix could be specified as:

$$\gamma_{ij} = \begin{cases} \gamma_{BB}/(N_k - 1) & \text{if both } i \text{ and } j \text{ are boys} \\ \gamma_{BG}/(N_k - 1) & \text{if } i \text{ is a boy and } j \text{ is a girl} \\ \gamma_{GB}/(N_k - 1) & \text{if } i \text{ is a girl and } j \text{ is a boy} \\ \gamma_{GG}/(N_k - 1) & \text{if both } i \text{ and } j \text{ are girls} \end{cases} \quad (4)$$

Thus, this specification distinguishes between within-gender and cross-gender interactions:  $\gamma_{BB}$  measures how a boy is affected by other boys, and  $\gamma_{BG}$  measures how a boy is affected by girls;  $\gamma_{GB}$  measures how a girl is affected by boys, and  $\gamma_{GG}$  measures how a girl is affected by other girls.

Table 8 presents the estimation results for this specification of the interaction matrix. Within-gender interactions for boys are particularly strong for time spent on jobs, for income from jobs as well as for income from parents, and for alcohol expenditures. For girls I find strong within-gender interactions for time spent on jobs, and for alcohol expenditures. Most of

the cross-gender interactions are asymmetric: the effect of girls on boys is generally larger than the effect of boys on girls.

A particularly remarkable pattern appears for alcohol expenditures: The cross-gender interactions are larger in magnitude than the within-gender interactions and the cross-gender interactions have different signs. Alcohol expenditures by boys are extremely responsive to alcohol expenditures by girls with a coefficient larger than 1. However, the more boys spend on alcohol, the lower the alcohol expenditures by girls. One story that would be consistent with this pattern is that boys pay for girls' drinks (note that the boys have more money available). In the model without social interactions I found a highly significant negative coefficient on the girl dummy variable, implying that girls spend about half as much on alcohol as do boys. This is a result that has been reported in the literature numerous times; *e.g.* Gaviria and Raphael (2001) and Dee and Evans (2003). In the present model, allowing for differences between within-gender and cross-gender interactions, the negative coefficient on the girl dummy (not reported) disappears (in fact it turns positive). This implies that in terms of the individual inclination to spend money on alcohol, there is no difference between boys and girls. (In fact, girls would spend slightly more than boys if each of them made his/her decision in isolation.) *It is the gender pattern of social interactions that causes boys to spend twice as much on alcohol as do girls, in the social equilibrium.*

The last column in table 8 reports  $p$ -values for testing  $\gamma_{BB} = \gamma_{BG} = \gamma_{GB} = \gamma_{GG}$ . At the 1 percent significance level, the hypothesis is rejected for time spent on jobs, income from jobs, and alcohol expenditures; at the 10 percent level it is also rejected for time spent on going out. Soetevent and Kooreman (2004) find similar results for smoking and truancy. For the other

types of behavior the hypothesis cannot be rejected. Thus jobs, income from jobs, going-out, smoking, drinking, and truancy show a coherent pattern of relatively strong and non-trivial social interactions within and between genders. This pattern also suggest that effective policy aimed at discouraging smoking, for example, should not merely focus on smoking but more generally on a life style of which smoking is only one aspect.

Recall that the results – while consistent with endogenous social interactions – do not formally establish their presence. However, many of the empirical results seem difficult to explain through another mechanism than endogenous social interaction, in particular the asymmetry in the boy-girls interactions.

Obviously, one could in principle refine the specification of the interaction matrix beyond the boy-girl distinction, for example on the basis of ethnicity, or by allowing the effects of younger and of older classmates to be different. The interactions could in principle be completely idiosyncratic. Some students – say leaders – might act completely independently of others. In the matrix  $\Gamma$  the row of such a student would merely contain zeros. The individual characteristics related to such independent behavior may be unobservable. This suggests a specification that allows for random variation in the parameters of the interaction matrix. Such refinements are left for future research.

Table 9 reviews some of the results from Soetevent and Kooreman (2004) for the discrete choice analogue of the model, with school specific fixed effects. The strongest endogenous social interactions are found for truancy, somewhat more moderate effects for smoking, cell phone ownership, and moped ownership, and weak effects for asking parent's permission for purchases. Thus endogenous social interactions are strong for behaviors closely

related to school (truancy), moderate for behavior partly related to school (smoking, cell phone and moped ownership), and weak for behavior far away from school. Again, within-gender interactions are stronger than interactions between genders. An exception here is moped ownership: for a boy the probability of moped ownership is strongly affected by moped ownership of other boys and of girls. Moped ownership of girls, on the other hand, is not affected by endogenous social interactions. Note that for cell phone ownership intra-girl interactions are stronger than intra-boy interactions.

In order to gain some insight in the magnitude of the social interaction effects implied by the estimated  $\gamma$ 's consider a reference class (largely based on median values of exogenous variables). This is a hypothetical MAVO class composed of 8 girls and 8 boys; all of them are aged 14, Dutch, non-protestant, non-catholic, and come from a two-parent household with a father working 36 hours per week and a mother working 16 hours per week. The estimated parameters imply that in equilibrium the expected number of truants is 3.14 (the probability of truancy is 0.191 for girls and 0.201 for boys). Now suppose that a surely truanting girl is added to this class (i.e. we add a girl with characteristics such that her probability of truancy is virtually equal to 1, irrespective of the behavior of others). Without social interaction effects, the expected fraction of truants would rise from 0.196 (3.14/16) to 0.244 (4.14/17), a 24 percent increase. Taking social interaction effects into account, the new equilibrium fraction of truants rises to 0.278 (4.73/17), an increase of 41 percent compared to the original level. The model also implies that a change in the value of an exogenous variable of only one of the pupils in principle affects the behavior of all pupils in class.

## 5 Theories related to teenage behavior

Which theories – or elements of theories – can help us to explain the empirical patterns shown here? The standard economic theory of investment in human capital is most closely related to "time spent on school" and "truancy". In the investment model students make a trade-off between higher effort now and higher expected future wages, or more current leisure and lower future wages. In a simple version of this model a student will spend more time on school the higher the future wages. The fact that students with higher school level (higher ability) spend more time on school and are less frequent truants is consistent with this prediction. Many of the other findings, however, cannot be explained easily along similar lines.

This section first looks at four categories of explanations for endogenous social interactions (or peer effects): *i*) interdependent resources, *ii*) interdependent preferences, *iii*) the theory of identity, and *iv*) endogenous social norm theory. This list is incomplete, and an eclectic selection from the literature. I am not aware of any behavioral theories that motivate the existence of exogenous social interactions.

### 5.1 Interdependent resources

A simple explanation for endogenous social interactions is the existence of a common resource constraint. In that case the share a student gets from a pie of given size depends negatively on the shares obtained by other students due to the adding-up constraint. In the behaviors considered in the previous sections, there is no such common resource constraint, except - perhaps – in the case of self-esteem. If self-esteem is purely a matter of ranking individuals within a school class, the total "amount" of self-esteem is given, and a negative endogenous social interaction effect will result. Note that the



coefficient for self-esteem was negative but insignificant.

## 5.2 Interdependent preferences

A straightforward way to incorporate social interactions in models of consumption and time allocation is by making the parameters of the utility function dependent on the choices of other individuals. An early example in which this idea is formalized is Pollak (1976).

Consider the following stripped-down version with only two consumers and two goods. Consumer 1 has income  $I_1$ ; he spends  $y_1$  on the first good and the remainder of his income,  $I_1 - y_1$ , on the second good. Consumer 2 has income  $I_2$ ; she spends  $y_2$  on the first good and the rest of her income,  $I_2 - y_2$ , on the second good. Let us specify the utility functions as

$$\begin{cases} U_1 &= \alpha \ln(y_1 - \gamma y_2) + (1 - \alpha) \ln(I_1 - y_1) \\ U_2 &= \alpha \ln(y_2 - \gamma y_1) + (1 - \alpha) \ln(I_2 - y_2) \end{cases} \quad (5)$$

with  $0 < \alpha < 1$  and  $\gamma$  not too large. Thus, the utility of consumer 1 depends not only on his own consumption of both goods, it also depends on how much the other person consumes of the first good. If the parameter  $\gamma$  is positive, there is a negative consumption externality. Maximization of (4a) by consumer 1 yields the optimal  $y_1$  as a function of  $y_2$ :

$$y_1 = \alpha I_1 + (1 - \alpha) \gamma y_2 \quad (6)$$

Similarly, I have for consumer 2

$$y_2 = \alpha I_2 + (1 - \alpha) \gamma y_1 \quad (7)$$

The Nash equilibrium corresponding to these reaction functions is

$$y_1^{\text{Nash}} = \frac{\alpha I_1 + (1 - \alpha) \gamma I_2}{1 - (1 - \alpha)^2 \gamma^2} \quad (8)$$

$$y_2^{\text{Nash}} = \frac{\alpha I_2 + (1 - \alpha) \gamma I_1}{1 - (1 - \alpha)^2 \gamma^2} \quad (9)$$

Without interdependent preferences ( $\gamma = 0$ ), both consumers spend a fraction  $\alpha$  of their income on the first good. Thus, with negative consumption externalities ( $\gamma > 0$ ), they spend more on the conspicuous (first) good, and less on the non-conspicuous (second) good, compared to the case where interdependent preferences are absent.

This simple model reveals another important assumption usually made implicitly in the empirical literature on social interactions, namely the assumption that observed behavior is a Nash equilibrium. Consider the case where preferences are still given by (4) with  $\gamma > 0$ , but now the consumers coordinate their decisions and jointly decide that each of them spends a fraction  $\alpha$  of income on the first good. It can be shown that in that case both consumers are better off than in the Nash equilibrium. Now suppose that a sample of consumer pairs would be available to estimate (7). If within each pair consumers cooperatively choose this Pareto improvement over the Nash equilibrium, the regression coefficient on the income of the other person would be insignificant. The example shows that it would be wrong to conclude that preferences are not interdependent.

The model also implies that if consumers actually behave Nash, then Pareto improvements can be induced by taxing the conspicuous (first) good and subsidizing the non-conspicuous (second) good. In the model there is no difference between physical negative externalities such as cigarette smoke or noise on the one hand, and negative externalities through human emotions such as envy on the other hand. Empirical evidence on this type of preference interdependence hence provides a welfare theoretic motivation for taxing such conspicuous goods, in addition to the more traditional arguments for taxation; see Kooreman and Schoonbeek (2004) for details and extensions.

### 5.3 Endogenous social norms

Another central concept in theoretical models on social interactions is a "norm". Essentially, a norm is an equilibrium phenomenon that arises because deviations from it are penalized and conformist behavior rewarded. Thus, social norms are endogenous. An interesting recent paper in which social norms are made explicit is Castronova (2003). Castronova presents an empirical study on social norms and sexual activity in US high schools. In his model, there is not only the binary choice of sexual activity; for those who choose not to be sexually active, there is the decision to be punitive or not with respect to those who are active, and with respect to those who do not punish. There is a cost of imposing stigma on others and there is a cost of having stigma imposed. In Castronova's model two types of Nash equilibria can occur, a punitive and a permissive equilibrium. In the punitive equilibrium, those who are sexually active are being stigmatized by those who are not; in the permissive equilibrium the fraction of sexually active students is larger than in the punitive one, and those who are not active do not stigmatize. The model can explain why behavior and norms may vary substantially across otherwise similar schools.

### 5.4 Theories of identity

Akerlof and Kranton (2002) have summarized and formalized some sociological theories on student behavior and schools. A central concept in these theories is identity. In the standard human capital model applied to the explanation of student effort, a student makes a trade-off between higher effort and higher future wages on the one hand, and more current leisure and lower future wages on the other hand (*cf.* Burgess et al. (2002)). In the theory of identity a student maximizes utility by making two choices. He chooses

his effort in school and he chooses his social category, *e.g.* jocks or nerds. In doing so he considers the match between his own actions and the ideal behavior of the chosen category. In this model a student may exert low levels of effort (or show other forms of vice behavior) in order to avoid a loss of self-image.

Is identity a just term that summarizes certain patterns of individual preferences and resources, or does it contain information in addition to that? For example, there may be no such thing as a group of nerds; those who are called nerds may just be those who are smarter and more academically inclined. The theory assumes that these groups exist, that they are clearly delineated, and that there is consensus among teenagers about who belongs to which group. Whether identity is a useful concept for empirical research is an open question.

## **5.5 Intergenerational interactions**

A final piece of explanation for teenage behavior I will discuss is related to intergenerational interactions. In addition to the effects of parental working time reported here, a number of authors report recent evidence on a direct, causal relationship between maternal and paternal employment and children's' cognitive and non-cognitive outcomes. An example in the non-cognitive domain is Anderson, Butcher, and Levine (2003), who provide evidence of a causal relationship between maternal employment and child overweight. These authors argue that time constraints of the mother affects her ability to monitor and supervise the child's nutritional intake and energy expenditure. (Again, the fact that the behavior of fathers is seemingly irrelevant might be related to the smaller variation in father's working hours in the data.)

Another notable paper in this respect is Haisken-DeNew and Bantle (2002). These authors show that – compared to the case with two non-smoking parents – the smoking probability of a teenager more than doubles if one of the parents smokes; if both parents smoke, the smoking probability triples (compared to the case with two non-smoking parents). As a consequence, policies targeted at reducing smoking by youths may fail or will be less effective if parents’ behavior is not taken into account.

Just as teenagers react to their parents, parents’ behavior is likely to be affected by their teenagers’ behavior. For example, a mother may decide to reduce her working hours if her child is a frequent truant; *cf.* Hill, Xeung, and Duncan (2001). Note that this effect implies a negative relationship between mother’s working hours and truancy. The fact that I find a positive relationship in the data suggests that the first effect – the effect of mother’s working hours on the teenager’s truancy – is the stronger one; see also Liu, Mroz and Van der Klaauw (2003), who focus on parental responses to a child’s cognitive outcomes.<sup>9</sup>

## 6 Conclusion

The simultaneous attention to a large number of aspects of teenage behavior – time, money, peers, and parents – in a single paper precludes an in-depth analysis of each of them. At the same time, the “broad” approach reveals interrelated patterns of behaviors which helps to understand the separate aspects. Smoking, drinking, truancy, going-out, and having a job appear to be closely related aspects of teenage behavior. One interpretation is that the concurrence of these behaviors define a “lifestyle” or “identity”, a latent variable on the basis of which social groups are formed and delineated. This finding suggests that a policy targeted at reducing teenage smoking, for ex-

ample, might be more effective if it addresses the life style of which smoking is one aspect, rather than by addressing smoking in isolation.

The explicit distinction between endogenous social interactions within genders and those between genders shows that the former is generally stronger than the latter, but also that there are notable exceptions. For alcohol, the gender based distinction of interaction types reveals an asymmetric gender pattern that causes boys to spend twice as much on alcohol as girls.

The empirical results are subject to qualifications regarding the identification assumptions. As emphasized by Manski and others, increasing the credibility of empirical results on social interactions is to a large extent a matter of the availability of appropriate data. Within the context of teenage behavior, we would like to be able to randomly reassign students from one class to another, or from one school to another (*cf.* Sacerdote (2001)). This may not be feasible in practice, but information from school splits and mergers, or information collected on students' school classes prior to and after moving to another city or region might come close to the ideal. Another potentially important piece of information usually not collected is on the timing of activities. For example, if teenagers appear to smoke primarily during weekdays and hardly during weekends, this would provide additional information on the nature of their social interactions. More generally, research on teenage behavior is likely to benefit from matched data on parents and children from the same sample of families.

As in the literature on intra-household decision making, employing non-cooperative or cooperative solution concepts is crucial in any analysis of social interactions. As yet, this issue has not been addressed in the empirical social interactions literature. The fact that classmates interact daily, often for many years, and become friends in many cases suggests that non-cooperative Nash

equilibria are not necessarily plausible.

## Notes

<sup>1</sup>In addition to the research motivations to be mentioned here, researchers can be motivated by personal circumstances. One researcher for whom this seems to be the case is *ESPE*'s first president, Bernard van Praag. When he became a father in the early seventies, he published on child benefits and the costs of child care (Van Praag, 1970); when his children were teenagers he published on teenage behavior (*cf.* Warnaar and Van Praag, 1997); his recent work is on ageing and pension systems (Van Praag and Cardoso, 2003). I readily acknowledge that my choice for the topic of the present paper is partly related to the fact that my own children are now teenagers.

<sup>2</sup>By first discussing data and then theories the paper deviates from the protocol usually adhered to in applied econometrics. However, "*there is ... an evolutionary interaction between theory and data that is central to the act of creation of empirical knowledge. The set of available models to analyze the data is never fixed in advance of looking at it. The data may suggest new models and the models may in turn suggest new data – or more careful examination of old data – in the light of new theoretical advances.*" (Quoted from Heckman and Neil (1996)).

<sup>3</sup>Variations in regional representation of schools and changes in the questions and coding make it difficult to compare sample averages across years. For example, for some time use categories, the 1990 survey had categorized responses for hours per day [0, 0-1, 0-2, ..., 9-10, >10], the 1992 and 1994 surveys had responses based on hours plus quarters per day, while the surveys as of 1996 had responses based on hours plus minutes per day. I opt for the 1992 data because this year had the most extensive questionnaire.

<sup>4</sup>This 'one-factor' correlation structure is equivalent with the formulation  $\epsilon_{ik} = \mu_k + \eta_{ik}$ , where  $\mu_k$  is a class-specific error term with variance  $\sigma_\mu^2$ , and  $\eta_{ik}$  an individual error term with variance  $\sigma_\eta^2$ ; all error terms uncorrelated. Then  $\rho = \sigma_\mu^2 / (\sigma_\mu^2 + \sigma_\eta^2)$  for two students from the same class and  $\rho = 0$  for two students from different classes.

<sup>5</sup>1 guilder  $\approx$  0.57 USD in 1992.

<sup>6</sup>All results reported on discrete choices (truancy, smoking, cell phone ownership, moped ownership, and asking permission for expenditures) are taken from Soetevent and Kooreman (2004).

<sup>7</sup>Identification is based on the problematic assumption that perceived behavior is not determined by actual behavior.

<sup>8</sup>In view of the average number of students within a class, one could in principle estimate the model separately for each class, and hence make all the parameters class-dependent. This makes the issue of identification different from the case considered by Moffitt (2001); in his case the assumed availability of only two observations per group precludes the estimation of group specific parameters.

<sup>9</sup>With respect to the literature on school quality and educational attainment recent research reports the usefulness of taking a broader view on child development than merely in terms of cognitive performance as measured by test scores. Kuhn and Weinberger (2002) and Heckman (2003) stress the importance of non-cognitive skills such as leadership and



personal discipline as predictors for labor market success later in life. Non-cognitive skills are even more likely to be important predictors when using a broader definition of success and performance, such as well-being or happiness.

## References

- Akerlof, G.A. and R.E. Kranton (2002), "Identity and Schooling: some Lessons for the Economics of Education", *Journal of Economic Literature*, vol. 40, pp. 1167-1201.
- Anderson, P., P. Levine, and K. Butcher (2003), "Maternal Employment and Overweight Children", *Journal of Health Economics*, forthcoming.
- Arcidiacono, Peter and Sean Nicholson (2002), *Peer Effects in Medical School*, NBER Working Paper #9025.
- Aronsson, T., Blomquist, S., and H. Sacklen (1999), "Identifying Interdependent Behavior in an Empirical Model of Labor Supply", *Journal of Applied Econometrics*, vol. 14, pp. 607-626.
- Castronova, E. (2003), "Social Norms and Sexual Activity in US High Schools", *Journal of Human Resources*, forthcoming.
- Dee, T.S. and W.N. Evans (2003), "Teen Drinking and Educational Attainment: Evidence from Two-Sample Instrumental Variables Estimates", *Journal of Labor Economics*, vol. 21, pp. 178-209.
- Duflo, Esther and Emmanuel Saez (2002), "Participation and investment decisions in a retirement plan: the influence of colleagues' choices", *Journal of Public Economics*, 85, pp. 121-148.
- Duflo, Esther and Emmanuel Saez (2003), "The Role of Information and Social Interactions in Retirement Plan Decisions: Evidence from a Randomized Experiment", *Quarterly Journal of Economics*, 118, pp. 815-842.
- Durlauf, S.N. and R.A. Moffitt (2003), "Introduction to the Special Issue: Empirical Analysis of Social Interactions", *Journal of Applied Econometrics*, vol. 18, pp. 499. pp. 815-842.
- Figlio, David and Jens Ludwig (2000), *Sex, Drugs, and Catholic Schools: Private Schooling and Non-Market Adolescent Behaviors*, NBER Working Paper #7990.
- Burgess, S., C. Propper, and K. Gardiner (2002), *The Economic Determinants of Truancy*, CASE discussion paper # 61, London School of Economics.
- Gaviria, A. and S. Raphael (2001), "School-Based Peer Effects and Juvenile Behavior", *Review of Economics and Statistics*, 83, pp. 257-268.
- Gruber, J. (2001), "Youth Smoking in the 1990's: Why Did It Rise and What Are the Long-Run Implications", *American Economic Review*, 91, pp. 85-90.

- Gruber, J. and J. Zinman (2001), "Youth Smoking in the US: Evidence and Implications", in J. Gruber (ed.), *Risky Behavior among Youth: an Economic Analysis*, University of Chicago Press, pp. 69-120.
- Haisken-DeNew, J.P. and C. Bantle (2002), *Smoke Signals: the Intergenerational Transmission of Smoking Behavior*, paper presented at the ESPE 2002 conference.
- Heckman, J.J. (2003), "The Supply Side of the Race Between Demand and Supply: Policies to Foster Skill in the Modern Economy", *De Economist*, vol. 151, pp. 1-34.
- Heckman, J.J. and D. Neal (1996), "Coleman's Contributions to Education: Theory, Research Styles and Empirical Research", in J. Clark (ed.): *James S. Coleman*, Falmer Press.
- Hill, M.S., W.-J.J. Yeung, and G.J. Duncan (2001), "Childhood family structure and young adult behaviors", *Journal of Population Economics*, vol. 14, pp. 271-299.
- Kandel, E. and E.P. Lazear (1992), "Peer Pressure and Partnerships", *Journal of Political Economy*, vol. 100, pp. 801-817.
- Kapteyn, A., S. Van de Geer, H. Van de Stadt and T. Wansbeek (1997), "Interdependent Preferences: An Econometric Analysis", *Journal of Applied Econometrics*, vol. 12, pp. 665-686.
- Kawaguchi, D. (2004), "Peer Effects on Substance Use among American Teenagers", *Journal of Population Economics*, vol. 17, pp. 351-367.
- Kooreman, P. (2000), "The Labeling Effect of a Child Benefit System", *American Economic Review*, vol. 90 (3), pp. 571-583.
- Kooreman, P. and L. Schoonbeek (2004), "Characterizing Pareto Improvements in an Interdependent Demand System", *Journal of Public Economic Theory*, vol. 6, pp. 427-443.
- Soetevent, A.R. and P. Kooreman, P. (2004), *A Discrete Choice Model with Social Interactions; with an Application to High School Teen Behavior*, working paper, University of Groningen.
- Kuhn, P. and C. Weinberger (2003), *Leadership skills and wages*, working paper University of California, Santa Barbara.
- Liu, H., T. Mroz, and W. van der Klaauw (2002), *Maternal Employment, Migration, and Child Development*, working paper, University of North Carolina at Chapel Hill.

- Manski, C.F. (1993), "Identification of Endogeneous Social Effects: The Reflection Problem", *The Review of Economic Studies*, 60, pp. 531-542.
- Manski, C.F. (2000), "Economic Analysis of Social Interactions", *Journal of Economic Perspectives*, 14, pp. 115-136.
- Pollak, R.A. (1976), "Interdependent Preferences", *American Economic Review*, 66, pp. 309-321.
- Sacerdote, B. (2001), "Peer Effects with Random Assignment: Result for Dartmouth Roommates", *Quarterly Journal of Economics*, 116, pp. 681-703.
- Thomas, D. (1990), "Intra-Household Resource Allocation; an Inferential Approach", *Journal of Human Resources*, vol. 24, pp. 635-664.
- Van Praag, B.M.S. (1970), "Kindercrèches: wie moet dat betalen?" ("Who should pay for child care facilities?"), *Economisch Statistische Berichten*, vol. 55, pp. 516-517.
- Van Praag, B.M.S. and P. Cardoso (2003), *How Sustainable are Old-age Pensions in a Shrinking Population with Endogenous Labour Supply?* CES Working Paper #861.
- Warnaar, M. and B.M.S. van Praag (1997), "How Dutch teenagers spend their money", *De Economist*, 145, pp. 367-397.
- Woittiez, I. and A. Kapteyn (1998), "Social Interactions and Habit Formation in a Model of Female Labour Supply", *Journal of Public Economics*, 70, pp. 185-205.

Table 1: Sample statistics

	min	mean	median	max	st. dev.	fraction zeros
<i>Time use</i>						
Sleep	35.0	59.1	59.5	91	6.9	0.000
Eating/personal care	0.0	9.5	7.0	35	5.9	0.026
School	14.0	38.2	38.0	70	6.7	0.000
Jobs	0.0	4.5	2.0	54	5.6	0.343
Household chores	0.0	4.4	3.5	35.8	5.3	0.252
Going out	0.0	4.8	4.0	40	5.2	0.297
TV/video/computer	0.0	20.6	17.5	119	15.14	0.029
Reading	0.0	5.8	3.5	35	5.6	0.145
Music	0.0	0.9	0.0	28	2.3	0.748
Sports	0.0	3.3	2.0	21	3.5	0.243
<i>Income</i>						
Income from parents	0.0	115.3	80	1200	114.6	0.070
Income from jobs	0.0	144.8	78	2567	198.4	0.248
<i>Expenditures and savings</i>						
Clothing and shoes	0.0	42.7	0.0	400	63.6	0.521
Alcohol	0.0	32.3	0.0	425.7	63.7	0.608
Savings	0.0	52.7	25	1000	83.9	0.190
<i>Self-esteem and well-being</i>						
Self-esteem	0.0	4.64	5	6	1.46	0.007
Well-being	1.0	5.03	5	6	0.87	0.000
<i>Student and family characteristics</i>						
Girl	0	0.481	0	1	0.500	
Age	11.0	15.6	16	23	1.63	
Non-Dutch	0	0.0410	0	1	0.198	
Single parent	0	0.0655	0	1	0.247	
Family size	1	4.28	4	4	0.992	
Urbanization	1	3.18	3	5	1.34	
School level 1 (MAVO)	0	0.292	0	1	0.455	
School level 2 (HAVO)	0	0.237	0	1	0.425	
School level 3 (VWO)	0	0.214	0	1	0.410	
Father's hours of work	0	41.2	40	70	14.2	
Mother's hours of work	0	18.7	18	70	15.9	
Father's college degree	0	0.221	0	1	0.415	
Mother's college degree	0	0.105	0	1	0.307	
Protestant	0	0.227	0	1	0.419	
Catholic	0	0.368	0	1	0.482	

Time use: hours spent per week during a normal school week

Income and expenditures: guilders per month (1 guilder $\approx$ 0.57 USD in 1992)

Self-esteem and well-being: see section 2 for definitions

Table 2: Reduced form regressions; time use ( $t$ -values in parentheses)

	Sleep	Eating/ personal care	School	Jobs	Household chores
Constant	81.2 (58.0)	5.8 (4.6)	42.6 (21.4)	-9.4 (-7.0)	8.0 (4.8)
Girl	0.40 (1.9)	0.68 (3.3)	1.05 (5.2)	-0.57 (-3.3)	0.62 (2.6)
Age	-1.29 (-17.7)	0.171 (2.5)	-0.370 (-3.2)	0.84 (12.3)	-0.13 (-1.5)
Non-Dutch	-2.00 (-3.4)	-0.62 (-1.0)	-0.64 (-1.0)	-1.55 (-3.4)	2.39 (4.0)
Single parent	-1.41 (-3.2)	0.96 (2.4)	-0.30 (-0.6)	-0.14 (-0.4)	1.91 (4.5)
Family size	-0.119 (-1.2)	0.130 (1.2)	-0.15 (-1.3)	0.25 (4.0)	0.21 (2.3)
Urbanization	-0.047 (-0.5)	0.062 (0.7)	-0.110 (-1.1)	0.00 (0.0)	0.10 (1.2)
School level 1 (MAVO)	-0.093 (-0.2)	-0.123 (-0.4)	1.96 (4.1)	-1.11 (-2.6)	-1.33 (-3.6)
School level 2 (HAVO)	-0.975 (-2.3)	0.416 (1.3)	3.13 (5.7)	-1.26 (-2.9)	-2.2 (-6.6)
School level 3 (VWO)	-1.60 (-3.7)	0.341 (1.0)	3.93 (7.2)	-1.93 (-3.7)	-3.0 (-7.7)
Fathers hours of work	-0.014 (-1.6)	0.006 (0.7)	-0.00 (-0.4)	0.015 (2.4)	-0.00 (-0.0)
Mothers hours of work	-0.009 (-1.2)	-0.004 (-0.6)	-0.011 (-1.5)	0.029 (4.9)	0.01 (1.8)
Father's college degree	-0.203 (-0.7)	0.244 (0.9)	-0.03 (-0.1)	-0.40 (-1.7)	0.00 (0.6)
Mother's college degree	-0.41 (-1.1)	-0.050 (-0.1)	0.06 (0.2)	-0.02 (-0.1)	0.20 (0.4)
Protestant	0.196 (0.6)	-0.464 (-1.6)	0.19 (0.6)	0.17 (0.8)	-0.19 (-0.8)
Catholic	0.212 (0.8)	-0.149 (-0.6)	0.41 (1.4)	0.16 (0.7)	-0.44 (-1.6)
$\hat{\sigma}^2$	39.5 (51.3)	31.3 (52.3)	40.5 (47.6)	25.1 (65.2)	25.9 (63.8)
$\hat{\rho}$	0.030 (2.8)	0.007 (1.0)	0.121 (8.7)	0.067 (7.1)	0.010 (0.8)

Reduced form regressions; time use (continued)

	Going out	TV/Video computer	Reading	Music	Sports
Constant	-10.0 (-7.8)	28.9 (11.7)	6.7 (4.5)	0.54 (0.9)	4.3 (4.8)
Girl	0.19 (1.1)	-2.14 (-5.7)	0.31 (1.4)	-0.15 (-1.6)	-0.93 (-7.2)
Age	0.93 (14.9)	-0.243 (-1.9)	0.01 (0.0)	0.00 (0.1)	-0.06 (-1.3)
Non-Dutch	-0.19 (-0.4)	2.09 (2.2)	0.05 (0.0)	-0.15 (-0.6)	-0.05 (-0.1)
Single parent	0.72 (2.2)	-0.73 (-0.9)	-0.02 (0.0)	-0.08 (-0.4)	-0.35 (-1.4)
Family size	-0.08 (-1.2)	-0.36 (-1.9)	0.16 (1.4)	0.01 (0.1)	-0.13 (-2.1)
Urbanization	0.067 (0.9)	-0.41 (-2.4)	-0.04 (-0.6)	0.04 (1.1)	-0.00 (-0.1)
School level 1 (MAVO)	-0.88 (-2.9)	-2.58 (-4.2)	-0.35 (-1.0)	0.06 (0.4)	0.40 (1.9)
School level 2 (HAVO)	-1.14 (-3.7)	-5.07 (-8.1)	-0.70 (-2.0)	0.30 (2.0)	0.62 (2.8)
School level 3 (VWO)	-1.94 (-6.0)	-7.2 (-10.3)	-0.80 (-3.3)	0.29 (2.0)	0.82 (3.7)
Fathers hours of work	0.023 (3.7)	0.02 (1.3)	-0.00 (-0.7)	-0.00 (-0.3)	0.016 (3.3)
Mothers hours of work	0.023 (4.4)	-0.00 (-0.0)	0.00 (0.3)	0.00 (0.1)	-0.00 (-0.8)
Father's college degree	0.142 (0.5)	-1.5 (-2.4)	0.16 (0.5)	0.44 (4.1)	-0.16 (-0.9)
Mother's college degree	0.291 (0.9)	-0.49 (-0.7)	0.33 (0.9)	-0.02 (-0.1)	0.46 (2.2)
Protestant	-0.414 (-1.8)	-1.21 (-2.5)	0.89 (3.3)	0.01 (0.9)	-0.17 (-1.0)
Catholic	-0.209 (-1.0)	-0.52 (-1.0)	-0.71 (-2.1)	0.07 (0.6)	0.11 (0.8)
$\hat{\sigma}^2$	21.3 (49.8)	112.3 (52.3)	29.1 (55.2)	5.4 (70.2)	11.6 (51.6)
$\hat{\rho}$	0.052 (3.8)	0.031 (3.1)	0.009 (0.9)	0.004 (0.5)	0.027 (3.2)

Table 3: Reduced form regressions; income ( $t$ -values in parentheses)

	Income from parents	Income from jobs
Constant	7.97 (0.2)	572.8 (-12.3)
Girl	-1.55 (-0.2)	-40.78 (-5.8)
Age	6.58 (3.0)	41.62 (17.2)
Non-Dutch	17.62 (1.2)	-23.34 (-1.4)
Single parent	49.90 (3.9)	13.42 (1.1)
Family size	-4.00 (-1.2)	12.86 (4.2)
Urbanization	1.58 (0.6)	4.94 (1.8)
School level 1 (MAVO)	-4.91 (-0.3)	-35.50 (-3.0)
School level 2 (HAVO)	-7.03 (-0.6)	-36.47 (-2.9)
School level 3 (VWO)	-18.6 (-1.5)	-56.01 (-4.6)
Fathers hours of work	0.24 (1.1)	0.71 (2.6)
Mothers hours of work	0.67 (3.3)	0.92 (4.7)
Father's college degree	18.1 (2.4)	-6.15 (-0.7)
Mother's college degree	1.73 (0.1)	-8.64 (-0.8)
Protestant	-15.62 (-1.6)	11.34 (1.1)
Catholic	-6.81 (-0.8)	2.49 (0.3)
$\hat{\sigma}^2$	11935 (42.3)	28856 (88.0)
$\hat{\rho}$	0.059 (2.5)	0.029 (3.1)



Table 4: Reduced form regressions; expenditures and savings ( $t$ -values in parentheses)

	Clothing and shoes	Cosmetics	Alcohol	Savings
Constant	-127.9 (-7.5)	-24.6 (-9.1)	-179.8 (-11.9)	-55.77 (-2.1)
Girl	16.90 (7.3)	6.37 (13.1)	-20.95 (-9.6)	-14.92 (-4.6)
Age	10.79 (13.2)	1.83 (12.1)	13.57 (17.6)	7.76 (5.7)
Non-Dutch	13.67 (3.4)	2.79 (3.1)	-13.04 (-2.2)	-16.11 (-1.2)
Single parent	7.92 (1.7)	1.86 (2.2)	1.04 (0.2)	-10.14 (-1.2)
Family size	0.09 (0.1)	0.10 (0.5)	1.06 (1.2)	0.94 (0.5)
Urbanization	-2.84 (-3.6)	-0.16 (-0.9)	0.96 (1.2)	0.05 (0.0)
School level 1 (MAVO)	-2.69 (-0.7)	-1.17 (-1.6)	-16.39 (-4.1)	-8.50 (-1.7)
School level 2 (HAVO)	3.24 (0.9)	-1.02 (-1.5)	-12.64 (-3.2)	-22.74 (-3.5)
School level 3 (VWO)	-2.27 (-0.5)	-2.07 (-3.2)	-20.96 (-4.5)	-22.98 (-3.8)
Father's hours of work	-0.02 (-0.2)	0.00 (0.2)	0.33 (4.3)	-0.01 (-0.1)
Mother's hours of work	0.31 (3.9)	0.04 (2.6)	0.18 (2.5)	0.31 (2.9)
Father's college degree	1.21 (0.4)	-0.60 (-0.7)	-0.52 (-0.2)	-3.12 (-0.6)
Mother's college degree	-2.07 (-0.5)	-0.45 (-0.4)	1.59 (0.4)	4.01 (0.6)
Protestant	-1.58 (-0.5)	-0.08 (-0.1)	-3.15 (-1.0)	3.09 (0.6)
Catholic	-5.09 (-1.9)	-1.02 (-1.6)	0.80 (0.3)	5.69 (1.4)
$\hat{\sigma}^2$	3264 (59.2)	131.6 (123.8)	3067 (76.5)	5877 (86.2)
$\hat{\rho}$	0.019 (2.1)	-0.011 (-1.6)	0.056 (5.4)	0.004 (0.5)

Table 5: Reduced form regressions; self-esteem and well-being ( $t$ -values in parentheses)

	Self-esteem	Well-being
Constant	4.72 (12.6)	6.06 (26.1)
Girl	-0.49 (-8.6)	-0.06 (-1.7)
Age	0.01 (0.4)	-0.05 (-4.3)
Non-Dutch	-0.26 (-1.7)	-0.10 (-1.1)
Single parent	-0.11 (-1.1)	-0.17 (-2.9)
Family size	-0.06 (-2.2)	-0.00 (-0.2)
Urbanization	-0.01 (-0.6)	-0.02 (-1.2)
School level 1 (MAVO)	0.17 (1.9)	-0.09 (-1.5)
School level 2 (HAVO)	0.26 (2.6)	-0.20 (-3.2)
School level 3 (VWO)	0.34 (3.3)	-0.11 (-1.8)
Father's hours of work	0.00 (1.1)	0.00 (0.2)
Mother's hours of work	-0.00 (-0.9)	0.00 (-0.4)
Father's college degree	0.17 (2.2)	0.00 (0.1)
Mother's college degree	0.05 (0.5)	-0.01 (-0.2)
Protestant	0.13 (1.5)	0.02 (0.4)
Catholic	0.09 (1.3)	0.05 (1.1)
$\hat{\sigma}^2$	2.02 (30.6)	0.70 (43.7)
$\hat{\rho}$	0.027 (2.4)	0.010 (1.2)

Table 6: Model with endogenous interactions ( $t$ -values in parentheses)

	$\hat{\gamma}$	$\hat{\rho}$	$\hat{\sigma}^2$
<i>Time use</i>			
Sleep	0.066 (0.7)	0.018 (0.9)	39.4 (51.5)
Eating/personal care		-	-
School	0.256 (2.3)	0.036 (1.1)	38.3 (37.8)
Jobs	0.237 (3.3)	0.015 (0.9)	24.8 (64.8)
Household chores	-	-	-
Going out	0.133 (1.7)	0.022 (1.1)	21.0 (51.5)
TV/video/computer	0.161 (1.9)	0.000 (0.0)	112.3 (52.3)
Reading	-	-	-
Sports	-	-	-
<i>Income</i>			
Income from parents	0.341 (5.8)	0.000 (-)	11618 (49.7)
Income from jobs	0.144 (1.9)	0.005 (0.4)	28818 (88.5)
<i>Expenditures and savings</i>			
Clothing and shoes	0.126 (1.7)	0.000 (-)	3241 (61.8)
Alcohol	0.309 (7.0)	0.000 (-)	2912 (78.6)
Savings	0.125 (1.3)	0.000 (-)	5755 (86.0)
<i>Self-esteem and well-being</i>			
Self-esteem	-0.133 (0.6)	0.056 (1.1)	2.04 (26.3)
Well-being	0.171 (2.7)	0.000 (-)	0.697 (43.1)

Table 7: Goodness-of-fit ( $R^2$ )

	Model without endogenous social interaction	Model with endogenous social interactions
<i>Time use</i>		
Sleep	0.166	0.168
Eating/personal care	0.090	-
School	0.092	0.140
Jobs	0.210	0.220
Household chores	0.061	-
Going out	0.199	0.209
TV/video/computer	0.510	0.515
Reading	0.070	-
Sports	0.050	0.058
<i>Income</i>		
Income from parents	0.092	0.116
Income from jobs	0.267	0.268
<i>Expenditures and savings</i>		
Clothing and shoes	0.194	0.199
Alcohol	0.245	0.283
Savings	0.166	0.183
<i>Self-esteem and well-being</i>		
Self-esteem	0.089	0.089
Well-being	0.083	0.085

Table 8: Intra-gender and cross-gender interactions (*t*-values in parentheses)

	$\hat{\gamma}$	$\hat{\gamma}_{BB}$	$\hat{\gamma}_{BG}$	$\hat{\gamma}_{GB}$	$\hat{\gamma}_{GG}$	<i>p</i> -value
<i>Time use</i>						
Sleep	0.066 (0.7)	-	-	-	-	-
Eating/personal care	-	-	-	-	-	-
School	0.256 (2.3)	0.217 (1.5)	0.234 (1.9)	0.249 (1.8)	0.257 (2.2)	0.545
Jobs	0.237 (3.3)	0.582 (11.6)	0.367 (2.7)	0.048 (0.5)	0.266 (3.0)	0.000
Household chores	-	-	-	-	-	-
Going out	0.133 (1.7)	0.149 (1.6)	0.348 (3.1)	-0.054 (-0.4)	0.169 (1.7)	0.054
TV/video/computer	0.161 (1.9)	0.185 (1.9)	0.200 (1.6)	0.155 (1.5)	0.100 (0.9)	0.730
Reading	-	-	-	-	-	-
Sports	0.145 (1.9)	0.194 (2.1)	0.023 (0.1)	0.200 (1.6)	0.006 (0.4)	0.195
Income from parents	0.341 (5.8)	0.392 (4.0)	0.254 (1.3)	0.147 (1.0)	0.135 (0.8)	0.421
Income from jobs	0.144 (1.9)	0.321 (3.9)	0.282 (2.3)	-0.022 (-0.2)	0.126 (1.0)	0.000
<i>Expenditures and savings</i>						
Clothing and shoes	0.126 (1.7)	0.142 (1.4)	-0.081 (-0.6)	0.339 (2.2)	0.190 (1.8)	0.116
Alcohol	0.309 (7.0)	0.271 (3.8)	1.337 (11.5)	-0.766 (-9.4)	0.468 (3.5)	0.000
Savings	0.125 (1.3)	-	-	-	-	-
<i>Self-esteem and well-being</i>						
Self-esteem	-0.133 (0.6)	-0.191 (-0.7)	-0.225 (-0.8)	-0.111 (-0.5)	-0.105 (-0.4)	0.761
Well-being	0.171 (2.7)	0.176 (1.7)	0.192 (1.8)	0.140 (1.8)	0.171 (2.2)	0.321

Table 9: Intra-gender and cross-gender interactions; discrete choices  
(*t*-values in parentheses)

	$\hat{\gamma}_{BB}$	$\hat{\gamma}_{BG}$	$\hat{\gamma}_{GB}$	$\hat{\gamma}_{GG}$
Truancy	0.880 (4.7)	0.533 (2.1)	0.569 (2.6)	0.765 (4.6)
Smoking	0.829 (6.8)	0.535 (3.5)	0.465 (2.9)	1.171 (10.3)
Cell phone	0.562 (5.1)	0.434 (2.8)	0.467 (2.7)	0.830 (8.2)
Moped	0.486 (2.4)	0.497 (2.0)	0.346 (1.1)	0.153 (0.6)
Asking permission	0.303 (2.1)	0.082 (0.5)	0.128 (0.8)	0.220 (2.0)

Source: Soetevent and Kooreman (2004)