

Exact computation of max weighted score estimators

Florios, Kostas; Skouras, Spyros

Postprint / Postprint

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

www.peerproject.eu

Empfohlene Zitierung / Suggested Citation:

Florios, K., & Skouras, S. (2008). Exact computation of max weighted score estimators. *Journal of Econometrics*, 146(1), 86-. <https://doi.org/10.1016/j.jeconom.2008.05.018>

Nutzungsbedingungen:

Dieser Text wird unter dem "PEER Licence Agreement zur Verfügung" gestellt. Nähere Auskünfte zum PEER-Projekt finden Sie hier: <http://www.peerproject.eu> Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under the "PEER Licence Agreement". For more information regarding the PEER-project see: <http://www.peerproject.eu> This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Accepted Manuscript

Exact computation of max weighted score estimators

Kostas Florios, Spyros Skouras

PII: S0304-4076(08)00077-8

DOI: [10.1016/j.jeconom.2008.05.018](https://doi.org/10.1016/j.jeconom.2008.05.018)

Reference: ECONOM 3047

To appear in: *Journal of Econometrics*

Received date: 16 August 2007

Revised date: 10 February 2008

Accepted date: 31 May 2008



Please cite this article as: Florios, K., Skouras, S., Exact computation of max weighted score estimators. *Journal of Econometrics* (2008), doi:[10.1016/j.jeconom.2008.05.018](https://doi.org/10.1016/j.jeconom.2008.05.018)

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

*Exact computation of max weighted score estimators**Kostas Florios¹ and Spyros Skouras²**¹Laboratory of Industrial and Energy Economics, School of Chemical Engineering,
National Technical University of Athens**²Department of International and European Economic Studies, Athens University of
Economics and Business*

Abstract: We show that exact computation of a family of ‘max weighted score’ estimators, including Manski’s max score estimator, can be achieved efficiently by reformulating them as mixed integer programs (MIP) with disjunctive constraints. The advantage of our MIP formulation is that estimates are exact and can be computed using widely available solvers in reasonable time. In a classic work-trip mode choice application our method delivers exact estimates that lead to a different economic interpretation of the data than previous heuristic estimates. In a small Monte Carlo study we find that our approach is computationally efficient for usual estimation problem sizes.

Keywords: maximum score, mixed integer programming, estimator computation, work-trip mode choice

JEL classification: C25; C61; C63 ; C87

1. Introduction

We are concerned with the computation of ‘max weighted score’ estimators that solve:

$$\max_{\beta} T^{-1} \sum_{t=1}^T w_t 1[(2y_t - 1)x_t \beta' > 0] \quad (1)$$

where y_t is a binary (0-1) dependent variable, x_t is an explanatory variable (a $p+1$ -dimensional row vector), w_t is a scalar weight on the objective function at t , $\beta = [\beta_0, \beta_1, \dots, \beta_p] \in \{-1, 1\} \times B$ is the parameter to be estimated (β_0 is normalized in this way to allow identification and is appropriately chosen from the $p+1$ parameters while B is typically compact) and $1[v]$ equals one if v is a true logical statement or zero otherwise. This objective function is closely related to certain objectives that are by now classic in the optimization literature, especially the ‘weighted max-sat’ objective (e.g., Borchers and Furman, 1999) and the max weighted feasible linear subsystems objective (e.g., Amaldi, Pfetsch and Trotter, 2003).

When $w_t = 1$ this estimator is the maximum score estimator of Manski (1975); when w_t is some arbitrary function of x_t it is Elliott and Lieli’s (2006) maximum utility estimator; and if $y_t = 1(v_t)$ and $w_t = |v_t|$ where v_t is an *observed* variable, it is an estimator for the sign of the regression of v on x , as shown by Skouras (2001). These estimators are of wide applicability, with the max score estimator having been used in empirical studies of, *inter alia*, work-trip mode choice (Horowitz, 1993), residential mobility determination (Bartik et al, 1992), idling of cement kilns (Das, 1991), entitlement of housing benefits (Blundell et al, 1988), forecasting of basketball games outcomes (Caudill, 2003), contingent valuation of forest resources (Li, 1996) and welfare

consequences to adult children due to taking care of their elderly parents (Kniesner et al, 2001). The estimator for the regression sign in the context of forecasting financial returns maximizes profits of a simple forecast-based investment strategy and is used for this purpose by traders (Skouras, 2001b).

The difficulty in computing these estimators arises because they are step functions of the parameters which furthermore will usually have a large number of local maxima (see Pinkse, 1993 and Skouras, 2003 for visualizations of such functions). This means standard optimization algorithms will perform poorly if they tend to get trapped in local maxima or may not be applicable, for example if they require analytical gradients. These observations are well known in the context of max score estimator computation which has attracted considerable interest (see Pinkse, 1993), and the treatment of the more general max weighted score estimators is very similar.

The most popular and probably the only widely available computational algorithm for computing max score estimators seems to be Manski and Thompson's (1986) 'great circle search algorithm' (GCS) which is implemented in LIMDEP. This is a heuristic algorithm (it does not guarantee a global optimum but is 'intuitively appealing') that has low computational complexity but does not guarantee global convergence (see Pinkse, 1993, p. 192) and as we shall see has poor performance in actual applications. As far as we are aware, it is the only algorithm that has been used in empirical applications of max score estimators (see Table C in our online appendix). An alternative heuristic is the smoothing procedure discussed in Skouras (2003) which however leads to exact solutions only under difficult to verify conditions. In the context of computing estimators, heuristic algorithms that guarantee only local optima are problematic because the statistical

properties of such procedures can differ from those of exact estimates in arbitrary ways, a point emphasized by Andrews (1997). Unfortunately, Andrews' method for dealing with this problem is inapplicable here due to the unusual nature of the weighted score objective function. The algorithm for exact computation of the maximum score estimator suggested by Pinkse (1993) compares scores at the intersections of the T hyperplanes x, β' (which partition the parameter space B into regions of constant score). This hyperplane intersection algorithm (HI) is too computationally expensive in realistic contexts: when $p=4$ and $T=500$, it requires the solution of approximately $2.5 \cdot 10^9$ systems of 4-by-4 linear equations.

We suggest a new method for exact computation of max weighted score estimators by reformulating them as a mixed integer programming problem (MIP). Our method has been designed to be more efficient than HI in realistic contexts, is based on MIP solvers available in many numerical mathematics packages and is therefore very easy to implement. Intuitively, the reason our estimator can achieve an *exact* optimum more efficiently than HI is because it determines parameter regions in which performance is low and avoids searching in those regions (see the example of Section 2). While MIP is still rarely used by economists, the fact that MIP reformulations of economic problems can be efficient was observed at least as early as Dantzig (1960); in an econometric context the only application we know of has been the recent work of Jouneau-Sion and Torres (2006) who use MIP to compute Dufour's (2006) maximized Monte Carlo test statistic.

The proposed MIP reformulation of max weighted score estimators is provided in the next section together with a very simple example that illustrates its effectiveness. In

Section 3.1 we use our software to re-compute the estimates in Horowitz's (1993) classic analysis of work-trip mode choice and arrive at an empirical model with very different economic interpretation and much larger score. In section 3.2 we provide a Monte Carlo comparison of the MIP algorithm with a Fortran implementation of Pinkse's HI algorithm and the LIMDEP implementation of Manski & Thompson's GCS algorithm. We make some concluding remarks in Section 4 and provide supporting documentation, auxiliary results, code for implementing our approach and a code manual in an online appendix at <http://liee.ntua.gr/kf/mws>. Our code can be linked to high-level languages or used as a stand-alone executable depending on the preferences of the user.

2. Mixed Integer Programming formulation of max weighted score estimators

For simplicity in what follows we will assume that the sign of β_0 is known on a priori grounds to be positive so $\beta_0=1$ (this negative case can be handled by multiplying x_0 by minus one). This is quite usual in practice and will often be the basis for choosing which variable will have the normalized parameter (i.e. be assigned to x_0). However, if such a priori information does not exist for any variable we would need to solve the max weighted score estimation problem twice, one for each possible value of β_0 .

It is straightforward to verify that an equivalent expression for the max weighted score estimator is:

$$\begin{aligned}
& \max_{\beta_1, \dots, \beta_p, z_1, \dots, z_T} \sum_{t=1}^T w_t z_t \quad s.t. \\
& (1-2y_t)x_t\beta' \leq [|x_{t0}| + \sup_{\beta_1, \dots, \beta_p} (|\sum_{i=1}^p x_{ti}\beta_i|)](1-z_t), \quad t=1, \dots, T, \\
& z_t \in \{0,1\}, t=1, \dots, T, \\
& [\beta_1, \dots, \beta_p] \in B,
\end{aligned} \tag{2}$$

where x_{ti} is the i 'th element of the row vector x_t ; remember that the zero'th element is the one for which the coefficient β_0 has been normalized to 1 which is emphasized by the formulation of the term in the brackets of the inequality.

This formulation introduces T new binary decision variables in addition to the parameter β (one z_t for each observation) that capture whether $\text{sgn}(x_t\beta')$ hits or misses $2y_t - 1$ at each t . The weighted hits are now maximised subject to T new *disjunctive constraints* on the estimator objective function. The computational difficulty of this problem depends on T , p and of course B , which controls the size of the search space, so as usual the range of plausible parameters needs to be chosen judiciously in the context of each application to be as narrow as possible.

Notice that this objective function is linear in decision variables (the step function has been eliminated) while some, but not all, decision variables are now constrained to be integer (in fact binary) variables. This linear mix of integer and continuous variables together with the disjunctive constraints represents a reformulation of our objective as a classic 'mixed integer linear program with disjunctive constraints' (Nemhauser and Wolsey 1999, p. 12).

Notice that if we relax the constraints on the z_t 's they become continuous variables in $[0,1]$ and the optimization problem becomes a standard linear programming problem which is easy to solve. The maximum of the relaxed problem will always be larger than

the maximum of the actual (more constrained) problem and this is exploited by standard MIP solvers such as branch-and-bound to easily eliminate directions of search in which the actual objective must be small since even the relaxed objective is small. In what follows we illustrate with a very simple example how the reformulated max weighted score estimation problem can be solved with a simple branch-and-bound algorithm, noting that the solvers used by our code are sophisticated refinements of the same approach.

Example: Consider data such that $x_0 = [2, -3, -1]'$, $x_1 = [1, 1, 1]'$, $y = [0, 1, 0]'$, $w = [1, 1, 1]'$ and $B = [-5, 5]$. In such a trivially small max score estimation problem a solution can be obtained effectively in a number of ways. The purpose of this example is only to illustrate the mixed integer linear program with disjunctive constraints formulation and how it can be solved; according to (2), max score estimation requires the solution of

$$\begin{aligned}
 & \max_{z_1, z_2, z_3, \beta_1} z_1 + z_2 + z_3 \text{ s.t.} \\
 & \beta_1 + 7z_1 \leq 5 \\
 & -\beta_1 + 8z_2 \leq 5 \\
 & \beta_1 + 2z_3 \leq 3 \\
 & z_1, z_2, z_3 \in \{0, 1\} \\
 & -5 \leq \beta_1 \leq 5
 \end{aligned} \tag{3}$$

This can be achieved for example by maximizing with respect to β_1 (a trivial exercise in linear programming) each of the eight possible combinations of values for $[z_1, z_2, z_3]$ and comparing these eight maxima to find the global maximum. A more intelligent branch-and-bound solution method with relaxations involves searching along the eight combinations by solving only four linear programs as follows:

Step 1: Solve the linear program (LP) produced by fixing z_1 to 1 and relaxing variables z_2, z_3 to take *any* value in $[0, 1]$. The solution is $[z_1, z_2, z_3, \beta_1] = [1, 0.375, 1, -2]$ with score

2.375. Notice that this is a solution that does not satisfy all the constraints of our desired objective since the constraint that the z 's should be binary have been relaxed.

Step 2: Now fix also z_2 to 1 allowing z_3 to be in $[0,1]$. This LP does not have a solution. Hence we can exclude both combinations $[z_1, z_2, z_3]=[1,1,1]$ or $[1,1,0]$ as potential solutions. This means there does not exist a parametrization which can score on both of the first two realizations.

Step 3: Fix $[z_1, z_2]=[1, 0]$ with z_3 in $[0,1]$. This LP is solved at $[1, 0, 1, -3.824]$ with score 2.

Step 4: In addition to the above, fix also z_3 to 1 to try a combination that satisfies all the constraints of the actual problem. A (non-unique) solution is at $[1,0,1,-2]$ with score 2.

Step 5: A solution with $[z_1, z_2, z_3]=[1,0,0]$ would have a score of one which is less than that at the parametrization of Step 4 so we do not need to solve the LP for this combination.

Step 6: Analogously to Step 1, we now fix z_1 at 0 and relax variables z_2, z_3 to take any value in $[0,1]$. This LP is solved at $[0,0.75,1,1]$ with a score of 1.75. This is lower than the solution of Step 4 which provides us with a score maximizing parameter $\beta_1=-2$ since imposing constraints on z_2, z_3 cannot possibly help.

We see that max score estimation was achieved using a branch-and-bound algorithm that required solving four trivial LP problems and just a few additional operations. In fact, MIP solvers will typically use heuristic refinements on branch-and-bound that can achieve large reductions in computation time.

3. Application: Computation of max score estimates

In order to assess the practical usefulness of the MIP formulation of the max weighted score estimator, we study in detail the computation underlying Horowitz's (1993) widely cited max score estimation of a work-trip mode choice model for Washington DC (also appearing in Horowitz (2004), Gozalo et al (2000) and McDonald (1996)). Our aim is to show that in a relevant problem the computational procedure we propose is practical, effective and a significant improvement over currently used methods. In our online appendix B we also provide results based on simulated data which agree qualitatively with the conclusions drawn in the context of this application.

The results reported below solve our MIP with disjunctive constraints formulation of max weighted score estimators by calling GAMS XPRESS - a widely available commercial solver for MIP problems (using a branch-and-cut technique, a variant of the branch-and-bound algorithm described in the example of section 2; see e.g. Williams (1985), pp.157-161). The code for obtaining these results has been included in the GAMS model library since Distribution 22.6 and is also available together with a similar open-source Fortran implementation in our online appendix. Commercial MIP solvers tend to be faster than the best open-source competitors, hence we have provided code that works with either. A convenient feature of either implementation is that they also provide the user with an upper bound on the underperformance relative to the global optimum should there be a computation time constraint that does not allow the computation of the global optimum.

3.1 Exact computation of Horowitz's (1993) work-trip mode choice model estimates

Briefly, for a sample of 842 persons in Washington DC in the late 1960's Horowitz (1993) modeled the 'work-trip mode choice' decision (automobile or other) for the daily trip from home to work. The explanatory variables were the number of cars owned by the traveler's household (CARS); the transit out-of-vehicle travel time minus automobile out-of-vehicle travel time (DOVTT in minutes); the transit in-vehicle travel time minus automobile in-vehicle travel time (DIVTT in minutes); and the transit fare minus automobile travel cost (DCOST in dollars). The coefficient chosen for the identification normalization (β_0) was that of DCOST (x_0) which was set to one. Table 1 reports published computed estimates from Horowitz (1993), estimates based on the LIMDEP implementation of the max score estimator (which uses essentially the same algorithm as Horowitz), and our own exact results. Horowitz reports computing these using code provided by Manski and Thompson that has been widely used, including Manski and Thompson (1986), Bartik et al (1992) and Das (1991). In order to impose a plausible constraint on B , we standardized variables to mean zero and unit standard deviation and (very conservatively) allowed each parameter to be in the range $[-10, 10]$ so that parameters were allowed to be an order of magnitude larger than the standard deviations in the raw data. Estimates obtained after this normalization also facilitate comparison of the relevance of each variable on the score.

Evidently, the published analysis would lead to the conclusion that DCOST and CARS are the most significant determinants of work-trip mode choice; we might easily arrive at the conclusion that DCOST is by far the most significant determinant of work-

trip mode choice since its parameter is twenty time larger than the next largest parameter (for CARS) with standardized data. This conclusion is also likely to emerge from LIMDEP estimates which however are even more puzzling in that they suggest the effect of CARS is *negative*; we have not attempted to evaluate statistical significance of the estimates because it is now known – as opposed to the time of publication of the original study - that standard errors for max score estimates are difficult to estimate, even with bootstrapping (Abrevaya et al, 2005). In contrast, our exact estimates indicate that CARS is by far the most important variable, and that probably the only factor of any relevance to work-trip mode choice is whether at least one car is owned - if CARS is one or larger the other variables must take on extremely negative values for the model to predict that the work-trip is not by automobile. This suggests the regression coefficients for DIVTT, DOVTT and DCOST should be treated as zero and interpreted in this way both economically and in any subsequent statistical analyses. This simple fact is obscured by computational inaccuracies in competing algorithms. The same inaccuracies probably corrupt published bootstrap estimates for standard errors and other statistics of max score estimates that are based on heuristic optimization procedures.

A more detailed auxiliary analysis confirmed this intuition: omitting DIVTT, DOVTT and DCOST led to a model in which car owners always use a car and the rest do not. This simple model reduces the score by only 11 hits and in particular the three additional variables lead to only 4 additional hits on car travelers who do not own a car and 7 car owners who do not travel by car. These additional hits are either due to a very negative value of DCOST ('other' is much cheaper than car transport) for car owners who in fact do not use their cars; or to a large value simultaneously in DIVTT, DOVTT and DCOST

for a non-car owner who nevertheless travels by car. Put simply, these variables only explain the behavior of a relatively small number of outliers.

3.2 Performance analysis of MIP approach

Having shown in the previous section that applications of the max score estimation should be based on exact estimates and that these can be achieved using a MIP approach, we now examine its time-performance. The computation of exact estimates using the GAMS XPRESS MIP solver (version 15.25) required significant computation time: on a standard (Windows XP) notebook with a 1.73 GHz Pentium M processor and 512MB RAM it took approximately 10.5 hours (37,516 CPU seconds to solve approximately 5 million linear programs). The total computation time depends on the quality of the solver used and we observed significant speed-ups with the CPLEX 11 solver and slow-down with the open-source SCIP solver. It is also crucial to note that in our experiments we observed that the exact maximum can often be obtained very quickly, within a few minutes, but what takes longer is for the algorithm to *prove* that the maximum found is indeed exact; this suggests that satisfactory results are obtainable even with severe computation constraints.

It is relevant to address the issue of how algorithm performance scales across the range of sizes of realistic applications and compare this with the performance of the only alternative available algorithm for computing exact estimates, i.e. Pinkse's (1993) HI algorithm, implemented by us in Fortran. As we discussed in the introduction for problems of the order of the work-trip mode choice application the HI algorithm will be unacceptably slow for most users with access to standard technology. Indeed, the exact HI Fortran algorithm took 658,001 CPU seconds to solve approximately 21 billion 4-by-4

linear systems – which is an order of magnitude slower than the MIP approach. In what follows, we evaluate the robustness of this result in realistically sized applications. We do not provide comparisons to results from GCS computations since we have already seen that such computations tend to be extremely inaccurate.

As it happens, the work-trip mode choice application of the previous subsection is average sized relative to the published literature. In order to gauge the scalability and randomness in performance across realistic applications, we performed a small Monte Carlo study the results of which are reported in Table 2. Specifically, ten samples of size 250, 500 and 1000 were randomly constructed by sampling without replacement from the data of the previous subsection. Since the original data contained only 842 observations, the Monte Carlo samples with 1000 observations were created by choosing 158 observations randomly without replacement from the original sample to occur twice in the Monte Carlo samples (together with the original 842 observations). We estimated linear models with two, three, four and five explanatory variables where the fifth explanatory variable was traveller income (INC, in thousands of dollars) - a variable not used in the original study. Thus in total, $3 \times 4 \times 10 = 120$ estimated models were computed.

From Table 2 two clear qualitative conclusions emerge: First, even for very small estimation problems the HI algorithm has no advantage over the MIP formulation. More importantly, it scales much worse than the MIP formulation and is probably impractical for most users if there are either more than 500 observations or more than three estimated parameters (the performance differential increases much more dramatically with the number of estimated parameters than sample size). Second, the MIP performance can

vary significantly from one data set to another, whereas the HI algorithm's performance is practically constant. This is not surprising since the MIP algorithm exploits sample-specific features of each optimisation to avoid unnecessary computations whereas the HI algorithm involves computations the number of which is invariant across problems of the same size.

Without emphasising quantitative differences in algorithm performance (since these can depend significantly on our software and hardware implementations which however are almost certainly advantageous for the HI algorithm) we can confirm what is expected from complexity considerations, i.e. that for medium sized problems the HI algorithm becomes impractical. But exact rather than heuristic computation of max score estimators is definitely necessary so computing max score estimators as MIP may be the only viable approach in realistically sized applications.

4. Concluding remarks

The purpose of this note is to communicate the observation that max weighted score estimators can be computed exactly using mixed integer programming methods and that this is practical in realistic applications. We apply our proposed MIP approach to max score estimation of a widely cited work-trip mode choice model and find our exact estimates lead to a different economic interpretation than published approximate estimates. In particular, for the vast majority, commuting by car occurs if and only if at least one car is owned, presumably because the decision to buy a car is very closely tied with the desire to drive to work; contrary to what was previously believed, travel costs and the total number of cars add no additional information. This illustrates the importance

of exact computation of estimators emphasized also by Andrews (1997) and suggests that ‘computation uncertainty’ (as an analog of estimation or model uncertainty) can be a major unquantified source of uncertainty about empirical results. The approach proposed here and the accompanying code we provide for computing max weighted score estimators should help eliminate this source of uncertainty in future studies.

We believe there is much scope for application of MIP in econometrics. In auxiliary results reported in online appendix E, we have shown that the sum of absolute deviations estimator of the ordered response model and the mode regression estimator of the censored regression model can also be treated as MIP problems; together with the application of MIP by Jouneau-Sion and Torres (2006) and our results here, this suggests that MIP may become a broadly useful tool for econometricians. The close connection between the weighted score and certain objectives studied by the Operations Research community is also an intriguing avenue for future research.

Acknowledgements

We are indebted to Joel Horowitz for providing the data used in this study as well as a number of very helpful comments and to George Mavrotas for extremely valuable advice.

The first author acknowledges financial support from the Greek State Scholarships Foundation (IKY), as well as from the NTUA Fundamental Research Senate Committee and the Program for Fundamental Research Reinforcement 'Leucippus'.

ACCEPTED MANUSCRIPT

References

- Abrevaya, J. and J. Huang, 2005, On the bootstrap of the maximum score estimator. *Econometrica* 73, 1175–1204.
- Amaldi, E., Pfetsch, M.E. and L.E. Trotter, Jr, 2003, On the maximum feasible subsystem problem, IISs and IIS-hypergraphs. *Mathematical Programming, Series A* 95, 533-554.
- Andrews, D.W.K., 1997, A stopping rule for the computation of generalized method of moments estimators. *Econometrica* 65, 913-932.
- Bartik, T.J., Butler, J.S. and J.T. Liu, 1992, Maximum score estimates of the determinants of residential mobility: implications for the value of residential attachment and neighborhood amenities. *Journal of Urban Economics* 32, 233-256.
- Blundell, R., Fry, V. and R. Walker, 1988, Modeling the take-up of means-tested benefits: the case of housing benefits in the United Kingdom. *Economic Journal* 98, 58-74.
- Borchers, B. and J. Furman, 1999, A two-phase exact algorithm for max-sat and weighted max-sat problems. *Journal of Combinatorial Optimization* 2, 299-306.
- Caudill, S.B., 2003, Predicting discrete outcomes with the maximum score estimator: the case of the NCAA men's basketball tournament. *International Journal of Forecasting* 19, 313-317.
- Dantzig, G.B., 1960, On the significance of solving linear programming problems with some integer variables. *Econometrica* 28, 30-44.
- Das, S., 1991, A semiparametric structural analysis of the idling of cement kilns. *Journal of Econometrics* 50, 235-256.

- Dufour, J.M., 2006, Monte Carlo tests with nuisance parameters: a general approach to finite-sample inference and nonstandard asymptotics. *Journal of Econometrics* 133, 443-477.
- Elliot, G. and R.P. Lieli, 2006, Predicting binary outcomes. University of California San Diego, Working Paper.
- Gozalo, P. and O. Linton, 2000, Local nonlinear least squares: using parametric information in nonparametric regression. *Journal of Econometrics* 99, 63-106.
- Horowitz, J.L., 1993, Semiparametric estimation of a work-trip mode choice model. *Journal of Econometrics* 58, 49-70.
- Horowitz, J.L., 2004, Semiparametric models, in: J.E. Gentle, W. Härdle and Y. Mori, (Eds.), *Handbook of Computational Statistics*, Springer.
- Jouneau-Sion, F. and O. Torres, 2006, MMC techniques for limited dependent variable models: implementation by the branch and bound algorithm. *Journal of Econometrics* 133, 479-512.
- Kniesner, T.J. and A.T. LoSasso, 2001, Intergenerational labor market and welfare consequences of poor health. Center for Policy Research Working paper No. 37, Maxwell School of Citizenship and Public Affairs, Syracuse University, New York.
- Li, C.Z., 1996, Semiparametric estimation of the binary choice model for contingent valuation. *Land Economics* 72, 426-473.
- Manski, C.F., 1975, Maximum score estimation of the stochastic utility model of choice. *Journal of Econometrics* 3, 205-228.
- Manski, C.F. and T.S. Thompson, 1986, Operational characteristics of maximum score estimation. *Journal of Econometrics* 32, 85-108.

- McDonald, J.B., 1996, An application and comparison of some flexible parametric and semi-parametric qualitative response models. *Economics Letters* 53, 145-152.
- Nemhauser, G.L. and L.A. Wolsey, 1999, *Integer and Combinatorial Optimization*, Wiley.
- Pinkse, C.A.P., 1993, On the computation of semiparametric estimates in limited dependent variable models. *Journal of Econometrics* 58, 185-205.
- Skouras, S., 2001, The sign of a mean regression: characterisation, estimation and applications. Available online at <http://ssrn.com/abstract=291922>.
- Skouras, S., 2001b, Financial returns and efficiency as seen by an artificial technical analyst. *Journal of Economic Dynamics and Control* 25, 213-244.
- Skouras, S., 2003, Global optimisation of the sum of random step functions. *Computational Statistics and Data Analysis* 42, 349-361.
- Williams, M.P., 1985, *Model building in Mathematical Programming*, Wiley, Belfast.

	INTCPT	CARS	DOVTT	DIVTT	DCOST	% Max score	Absolute Max score
MIP (raw)	-110.08	170.67	3.630	0.8477	1.0	0.9086	765
MIP (norm)	5.122	3.916	0.962	0.401	1.0	0.9086	765
GCS (raw)	47.78	4.152	0.3785	-0.7475	-1.0	0.7850	661
GCS (norm)	1.5558	0.09526	0.1003	-0.3536	-1.0	0.7850	661
HOR (raw)	-1.6466	2.2520	0.0411	0.0110	1.0	0.5736	483
	(0.1374)	(0.1480)	(0.0294)	(0.0106)			
HOR (norm)	-0.27619	0.05167	0.01089	0.00520	1.0	0.5736	483

Table 1: Parameter and max score results for Horowitz (1993) work-trip mode choice data by exact mixed integer program (MIP) and heuristic LIMDEP implementation of the great circle search (GCS) method together with results reported in Horowitz (1993) as computed by an alternative implementation of GCS. The parameters are reported for raw and standardised (to mean zero and unit standard deviation) data. In an auxiliary analysis we exhaustively searched the parameter space (for the standardized regressors) and found that the parameter region in which the exact maximum and near maxima are achieved is both narrow and unique.

T		p							
		2		3		4		5	
		HI	MIP	HI	MIP	HI	MIP	HI	MIP
250	Mean	0.244	0.24	22.22	1.6	1566	10	87806	89
	Std.Dev.	0.003	0.06	0.14	0.9	9	6	243	79
	Median	0.244	0.23	22.19	1.3	1564	10	87695	58
	Min	0.240	0.16	22.02	0.9	1559	3	87580	17
	Max	0.251	0.36	22.42	3.7	1583	26	88330	244
500	Mean	1.960	1.4	350.4	37	48536	920		
	Std.Dev.	0.022	0.3	2.3	15	202	666		
	Median	1.958	1.3	350.5	33	48559	761		
	Min	1.924	1.0	345.7	22	48154	190		
	Max	2.003	1.8	354.0	66	48816	2269		
1000	Mean	16.06	12	5729	4339				
	Std.Dev.	0.14	4	42	1885				
	Median	16.04	12	5706	4563				
	Min	15.77	7	5690	1989				
	Max	16.35	17	5796	6584				

Table 2: CPU time in seconds for HI and MIP algorithms using simulation data. The MIP runs were executed on a notebook with a Pentium M processor of 1.73 GHz and 512MB RAM running Windows XP using GAMS XPRESS software. The HI runs were executed on a workstation with two dual core AMD Opteron 880 processors (only one processor was used for this run) and 8GB RAM running SuSe Linux64 using g77 open source FORTRAN77 compiler. A graphical representation of this table is available in online appendix D.

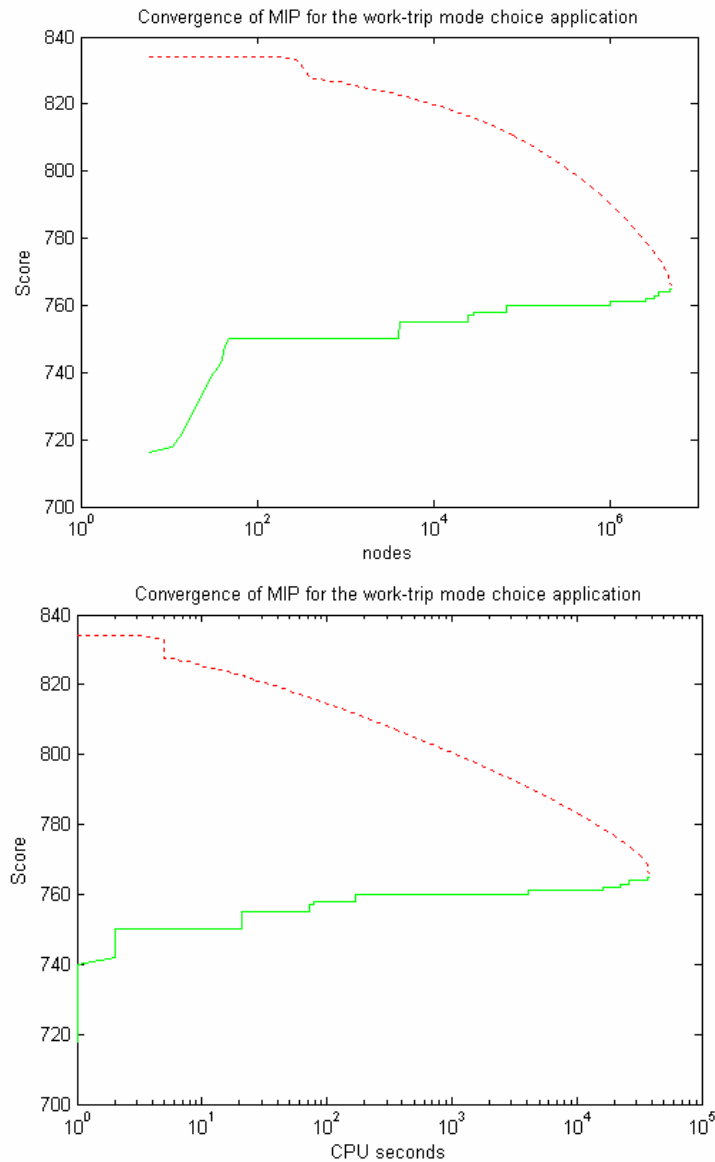


Figure 1: Convergence of the MIP approach applied to the work-trip mode choice max score estimator. The dashed line represents an upper bound on the max score and the solid line represents the computed maximum score as a function of time (nodes in the first panel and CPU secs in the second). In total, 4,955,300 nodes are solved and account for 37516 CPU sec. The run was executed on a notebook with a Pentium M technology processor of 1.73 GHz and 512MB RAM running Windows XP using GAMS XPRESS software.