Consistent subsets: Computing the Houtman-Maks index in Stata

PRELIMINARY DRAFT

Marcos Demetry
Linnaeus University
Växjö, Sweden
marcos.demetry@lnu.se

Per Hjertstrand
Research Institute of Industrial Economics
Stockholm, Sweden
per.hjertstrand@ifn.se

This version: June 19, 2022

Abstract.   The Houtman-Maks index is a measure of the size of a violation of utility maximizing (i.e., rational) behavior. This note introduces the *Stata* command `hmindex`, which calculates the Houtman-Maks index for a data set of prices and observed choices of a consumer. The command is illustrated with an empirical application.

Keywords: hmindex, Houtman-Maks index, revealed preference, WGARP, WARP

1 Introduction

The hypothesis that economic agents choose commodity bundles by maximizing a utility function subject to a budget constraint forms the core of neoclassical economics. Revealed preference is an efficient tool to test whether consumer choice data satisfies utility maximizing (i.e., rational) behavior.[1] When these data violates utility maximization, it is often desirable to know the "size" of the violation. Houtman and Maks (1985) propose to measure the degree of inconsistency as the maximal number of observations in the observed sample consistent with rational choice. This measure (the HM-index) is calculated as the maximal subset of observations consistent with some revealed preference axiom.[2]

This note introduces the *Stata* command `hmindex`, which calculates the HM-index for a data set of prices and observed choice quantities sampled from an individual consumer (or a cross-section of individuals). The `hmindex`-command implements a combinatorial algorithm proposed by Gross and Kaiser (1996) to calculate the HM-index for the weak generalized axiom of revealed preference (WGARP) and the weak axiom of revealed

---

1. See Demetry et al. (2022) for a brief introduction to empirical revealed preference theory.
2. There exist several other goodness-of-fit measures for revealed preference tests. One of the most prominent is the Afriat efficiency index (AEI) which is implemented in the `aei` command documented in Demetry et al. (2022). The HM-index is a more disaggregated goodness-of-fit measure than the AEI since it gives a binary response to whether or not a specific observation is included in the maximal subset consistent with revealed preference.

preference (WARP), which we refer to as the GK-algorithm.

When the data consists of two goods, WGARP (WARP) is a necessary and sufficient condition for the data to be (strictly) rationalized by a continuous, strictly increasing, and (strictly) concave utility function. Thus, in such case, the HM-index gives the maximal subset of observations that is consistent with rational choice. For data sets with more than two goods, WGARP/WARP are only necessary conditions for rationality, since they do not account for transitive binary relations. Thus, for such data, the HM-index gives an upper bound on the maximal subset of observations consistent with rationality. However, in many multi-dimensional data sets violations of rationality seem almost invariably also WGARP violations. That is, higher order insensitivities in the data detected by more stringent tests are often associated with pairwise inconsistencies that are detected by WGARP (Alston and Chalfant 1991). If so, the HM-index would give the correct maximal subset consistent with rationality.

A feature of our command is that it allow the user to calculate the distribution of the HM-index over uniformly random distributed data. We illustrate the command on experimental data collected by Choi et al. (2007).

2 The GK-algorithm to calculate the HM-index

Suppose there are T observations of the prices and quantities of $K \geq 2$ goods. At observation $t = 1, \ldots, T$, the prices and quantities are denoted by $\mathbf{p}^t = (p_1^t, \ldots, p_K^t)$ and $\mathbf{x}^t = (x_1^t, \ldots, x_K^t)$, respectively. We assume that all prices are strictly positive, and that all quantities are non-negative (i.e., some but not all quantities at any given observation may be equal to zero). The data set $(\mathbf{p}^t, \mathbf{x}^t)_{t=1,\ldots,T}$ usually describe a single consumer that is observed over time, but can also describe a cross-section of consumers at a fixed time period.[3]

For any pair of observations (t, s), we say that $\mathbf{x}^t$ is *directly revealed preferred* to $\mathbf{x}^s$, written $\mathbf{x}^t R^D \mathbf{x}^s$, if $\mathbf{p}^t \cdot \mathbf{x}^t \geq \mathbf{p}^t \cdot \mathbf{x}^s$. This means that $\mathbf{x}^t$ is chosen even though the cost of the bundle $\mathbf{x}^s$ (at prices $\mathbf{p}^t$) does not exceed $\mathbf{p}^t \cdot \mathbf{x}^t$. For all $\mathbf{x} \in \mathbb{R}_+^K$ and any $t = 1, \ldots, T$ such that $\mathbf{p}^t \cdot \mathbf{x}^t \geq \mathbf{p}^t \cdot \mathbf{x}$, the data $(\mathbf{p}^t, \mathbf{x}^t)_{t=1,\ldots,T}$ is *rationalized* by a utility function u if $u(\mathbf{x}^t) \geq u(\mathbf{x})$, and *strictly rationalized* if $u(\mathbf{x}^t) > u(\mathbf{x})$ whenever $\mathbf{x} \neq \mathbf{x}^t$.[4]

A data set $(\mathbf{p}^t, \mathbf{x}^t)_{t=1,\ldots,T}$ satisfies the *weak generalized axiom of revealed preference*, abbreviated WGARP, if $\mathbf{x}^t R^D \mathbf{x}^s$ implies $\mathbf{p}^s \cdot \mathbf{x}^s \leq \mathbf{p}^s \cdot \mathbf{x}^t$.

When the data consists of only two goods, it is well-known that WGARP is a necessary and sufficient condition for a data set to be rationalized by a continuous, strictly increasing, and concave utility function (See Varian (1982) in combination with Banerjee and Murphy (2006)).

---

3. If the data forms a cross-section then the problem of calculating the maximal subset of the data consistent with revealed preference can be thought of as a problem of finding the maximal number of consumers that share a common utility function (i.e., the same preferences).

4. Demetry et al. (2022) discuss the economic interpretation of rationalizability.

When the data consists of more than two goods, Aguiar et al. (2020) shows that WGARP (WARP) is a necessary and sufficient condition for a data set to be (strictly) rationalized by a continuous, strictly increasing, piecewise (strictly) concave, and skew-symmetric preference function.[5]

A data set $(\mathbf{p}^t, \mathbf{x}^t)_{t=1,\ldots,T}$ satisfies the *weak axiom of revealed preference*, abbreviated WARP, if $\mathbf{x}^t R^D \mathbf{x}^s$ implies $\mathbf{p}^s \cdot \mathbf{x}^s < \mathbf{p}^s \cdot \mathbf{x}^t$ whenever $x^t \neq x^s$.

In the two-dimensional case, WARP is a necessary and sufficient condition for a data set to be *strictly* rationalized by a continuous, strictly increasing, and *strictly* concave utility function (See Matzkin and Richter (1991) in combination with Rose (1958)). The difference between WGARP and WARP in economic terms is that WGARP allows for indifference between bundles. In other words, while WARP is only consistent with demand functions, WGARP can also accommodate demand correspondences.

The HM-index is defined as the maximal subset of observations from the data $(\mathbf{p}^t, \mathbf{x}^t)_{t=1,\ldots,T}$ such that WGARP/WARP holds.[6] All observations that are removed from the data form the *violator set* (VS), while the remaining observations satisfying WGARP/WARP form the *consistent set* (CS). The value of the HM-index is usually presented as either: (i) the *number* of observations in CS, i.e., the maximal number of observations satisfying WGARP/WARP, or (ii) the *fraction* of observations in CS, i.e., the number of observations in CS divided by the total number of observations in the data.

To calculate the HM-index, Gross and Kaiser (1996) use a graph-theoretic approach.[7] Every observation is interpreted as a node of a graph. If observations s and t form a violation of WGARP, then the nodes for s and t are adjacent. The degree of a node t, $degr(t)$, is the number of nodes to which it is adjacent. Define At as the set of nodes adjacent to node t, and $1At$ as the set of nodes which are adjacent to t with degree 1. The GK-algorithm consists of two parts:

Step 1: Whenever $degr(t) = \max_{s=1,\ldots,T} degr(s)$ and $degr(r) < degr(t)$ for all $r \in At$, then remove observation t.

Step 2: Whenever $degr(t) = degr(i) = \max_{s=1,\ldots,T} degr(s)$ and $i \in At$, then (a) if $1At = \varnothing$, remove observation t; (b) if $1Ai = \varnothing$, remove observation i, and (c) if $1At = 1Ai = \varnothing$, remove either t or i.

Steps 1 and 2 are repeated sequentially until no additional observation is removed. Heufer and Hjertstrand (2015) suggested to use the GK-algorithm for WGARP. Gross and Kaiser (1996) originally suggested to use the algorithm for WARP, which only

---

5. See Aguiar et al. (2020) for the definition of a preference function and the relevant properties pertaining to preference functions.
6. Strictly speaking, the HM-index is defined as the maximal subset of observations such that some revealed preference axiom holds. But since we are only concerned with WGARP/WARP, we only consider the more narrow definition stated here.
7. The standard way of calculating the HM-index in empirical applications have been to iteratively delete observations and test for revealed preference. However, even for relatively small data sets, this can be very computationally demanding and sometimes even practically unfeasible.

requires redefining adjacency. If two nodes, t and s, are defined as adjacent whenever t and s form a violation of WARP, then the GK-algorithm will provide the set of indices consistent with WARP.

Gross and Kaiser (1996) point out that the algorithm may remove additional observations, in which case, it only produce a lower bound and not an exact solution to the HM-index. However, they also argue, based on experimental evidence, that this should occur very rarely.

As a benchmark to rational choice, empirical applications of the HM-index using experimental data sometime compare the distribution of the HM-index across all subjects with the distribution of the HM-index under the assumption of uniformly random consumption behavior. The `hmindex`-command allow the user to calculate the empirical distribution of the HM-index under this type of irrational consumer behavior, by drawing uniformly random data as explained in Section 2.4 of Demetry et al. (2022) and, for every simulated data set, calculate the HM-index. The distribution can then be visualized by plotting the values of the HM-index in a kernel density plot (See Figure 2 in Section 4 for such an application to experimental data).

3   The hmindex command

The `hmindex`-command is available on SSC and can be installed by entering 'SSC install hmindex' at the Stata command prompt.

Syntax

hmindex, `price`(*string*) `quantity`(*string*) $\big[$ <u>axiom</u>(*string*) <u>distribution</u>
     <u>sim</u>ulations(#) <u>seed</u>(#) $\big]$

`price`(*string*) specifies a $T \times K$ price matrix, where each row corresponds to an observation t and the two columns correspond to the two goods. All prices are required to be strictly positive. If any of the elements in the price matrix are non-positive (or if the price and quantity matrices have different dimensions), the command returns an error message.

`quantity`(*string*) specifies a $T \times K$ quantity matrix, where each row corresponds to an observation t and the two columns correspond to the two goods. All quantities are required to be non-negative. Some (but not all) quantities at a given observation may be equal to zero. If the quantity matrix violates these conditions (or if the price and quantity matrices have different dimensions), the command returns an error message.

Options

<u>axiom</u>(*string*) specifies the axiom(s) that the user would like to apply. The default
option is axiom(WGARP). The user can apply WARP by specifying axiom(WARP).
The user may also apply both axioms simultaneously by specifying axiom(all).

<u>distribution</u> specifies whether the user would like to calculate the empirical distribu-
tion of the HM-index under uniformly random consumption behavior. The simulated
uniformly random data are calculated as explained in Sections 2.4 and 3.3 of Deme-
try et al. (2022). The user can set the number of simulations and the random seed
in the simulation of the uniformly random data (See the next two options). The
default option is that distribution is *not* specified.

<u>simulations</u>(#) specifies the number of repetitions of the simulated uniformly random
data sets. The default number of repetitions is simulations(1000). This option is
only useful in combination with the distribution option.

<u>seed</u>(#) specifies the random seed in generation of uniformly random data sets (See
Sections 2.3 and 3.3 of Demetry et al. (2022) for a detailed explanation). The default
random seed is seed(12345). This option is only useful in combination with the
distribution option.

Stored results

hmindex stores the following in r():

Scalars
r(OBS)	number of observations
r(GOODS)	the number of goods
r(HM_NUM_*axiom*)	maximal number of observations that satisfy the axiom. Given by #HM in the output
r(HM_FRAC_*axiom*)	the maximal fraction of observations that satisfy the axiom. Given by %HM in the output and calculated as %HM = #HM/T

Macros
r(AXIOM)	axiom or axioms being tested

Matrices
r(CS_price_*axiom*)	returns the price matrix in CS for the specified axiom, i.e., the price data corresponding to the goods in the consistent set
r(CS_quantity_*axiom*)	returns the quantity matrix in CS for the specified axiom, i.e., the quantity data corresponding to the goods in the consistent set
r(VS_price_*axiom*)	returns the price matrix in VS for the specified axiom,, i.e., the price data corresponding to the goods in the violator set
r(VS_quantity_*axiom*)	returns the quantity matrix in VS for the specified axiom,, i.e., the quantity data corresponding to the goods in the violator set.
r(SUMSTATS_*axiom*)	summary statistics for random data: number and fraction of observations in consistent set
r(SIMRESULTS_*axiom*)	number and fraction of observations in consistent set for every simulated uniformly random dataset

Examples

The following two examples illustrate `hmindex` using a dataset of 20 observations on the prices and quantities of two goods. The consumed quantities of goods 1 and 2 are `x1` and `x2`, which form the quantity matrix `X=(x1,x2)`. The corresponding price matrix with prices `p1` and `p2` of the two goods is `P=(p1,p2)`. The first example runs `hmindex` using its default options, i.e., for WGARP.

```
. use "hmindex_example_data.dta", clear
. mkmat x1 x2, matrix(X)
. mkmat p1 p2, matrix(P)
. hmindex, price(P) quantity(X)
    Number of obs            =        20
    Number of goods          =         2
```

Axiom	#HM	%HM
WGARP	15	.75

The second example runs `hmindex` for WARP with the `distribution` option using 1,000 simulations.

```
. hmindex, price(P) quantity(X) ax(warp) dist sim(1000)
    Number of obs            =        20
    Number of goods          =         2
    Simulations              =      1000
```

Axiom	#HM	%HM
WARP	15	.75

Summary statistics for simulations:

WARP	#HM	%HM
Mean	15.809	.79045
Std. Dev.	1.510232	.0755116
Min	11	.55
Q1	15	.75
Median	16	.8
Q3	17	.85
Max	20	1

4   Empirical illustration

We applied the `hmindex`-command to experimental data collected by Choi et al. (2007). These data consists of portfolio choice allocations in a two-dimensional setting from 93 experimental subjects over 50 decision rounds, i.e., $T = 50$. Each subject split her

budget between two Arrow-Debreu securities, with each security paying 1 token if the corresponding state was realized, and 0 otherwise. The experiment consisted of two treatments. In the first (symmetric) treatment with 47 subjects, each state of the world occurred with probability 1/2, which were objectively known to the subjects. The second (asymmetric) treatment was applied to 46 subjects, where the subjects faced states occurring with objective probabilities 1/3 and 2/3. All state prices were chosen at random and varied across all decision rounds and subjects.
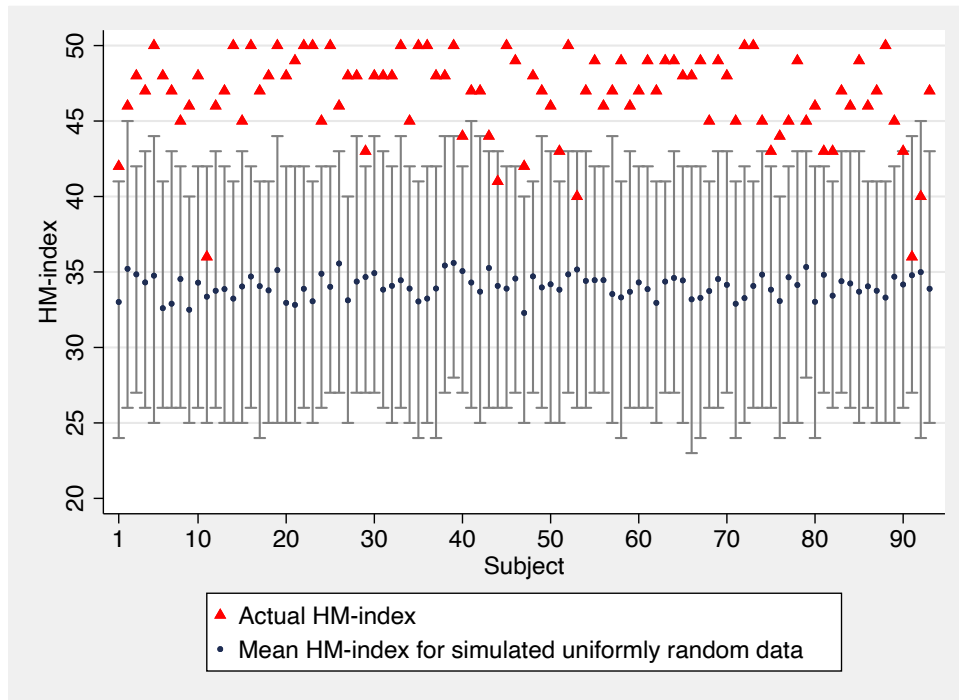

Figure 1: Calculated HM-indices from actual choices (red-colored diamond markers) and from simulated uniformly random data (black dots refer to the mean HM-index over the simulated data)

Choi et al. (2007) reported the HM-index for all but six subjects which they were unable to find an optimal solution for. Using the `hmindex`-command, we calculated the HM-index for everyone of the 93 subjects including the six unreported subjects in Choi et al. (2007). For every subject, the following code calculates (and saves to the matrix `results`) the actual HM-index and the mean, minimum and maximum HM-indices over all simulations in the uniformly random data.

```
matrix results = J(93,4,0)
matrix colnames results = "HM" "Mean" "Min" "Max"

forvalues subject = 1/93 {
```

```
quietly hmindex, price(P`subject´) quantity(X`subject´) dist sim(1000)

matrix sumstats = r(SUMSTATS_WGARP)

matrix results[`subject´, 1] = r(HM_NUM_WGARP)        /* Actual HM */
matrix results[`subject´, 2] = sumstats[1, 1]         /* Mean */
matrix results[`subject´, 3] = sumstats[3, 1]         /* Min */
matrix results[`subject´, 4] = sumstats[7, 1]         /* Max */

    }
```

Figure 1 presents these results.[8] The red-colored diamond markers represent the calculated HM-index for each subject. The connected intervals give the maximum and minimum values of the HM-index calculated from the simulated uniformly random data, while the black dots give the mean HM-index calculated over all simulations. We find that 77 subjects have a HM-index of 45 or higher (i.e., %HM $\geq$ 0.9), 14 subjects have values between 40 and 44 (i.e., $0.8 \leq$ %HM < 0.9), and 2 subjects have values below 40 (i.e., %HM < 0.8). Moreover, of all 93 subjects, the HM-indices from 6 of them are lower than the maximum HM-indices calculated from the simulated uniformly random data sets, which may cast doubt on whether the actual choices from these subjects are rational.

In Figure 2, we plot the kernel cumulative distribution functions (left) and probability distribution functions (right) of the HM-indices calculated from the subjects' actual choices (93 observations, solid line) and the HM-indices from the simulated uniformly random data ($93 \times 1,000 = 93,000$ observations, dashed line). A one-sided Kolmogorov-Smirnov test of the null hypothesis that the distribution of actual choices first-order stochastically dominates the distribution of uniformly random choices cannot be rejected at any conventional nominal significance level (p-value < 0.0000).[9] Overall, these results show that the actual choices are more consistent with rational choice than uniformly random choices.

Finally, we note that the `hmindex`-command is very quick: It found a solution for every subject in at most 0.036 seconds. The mean running time over all subjects was 0.0159 seconds with a standard deviation of 0.0063 seconds.

5   Acknowledgments

6   References

Aguiar, V. H., P. Hjertstrand, and R. Serrano. 2020. A rationalization of the weak axiom of revealed preference. Bravo Working Paper 2020-016, Brown University and IFN Working Paper No. 1321.

---

8. The help-file accompanying the `hmindex`-command contains the code to generate Figure 1.

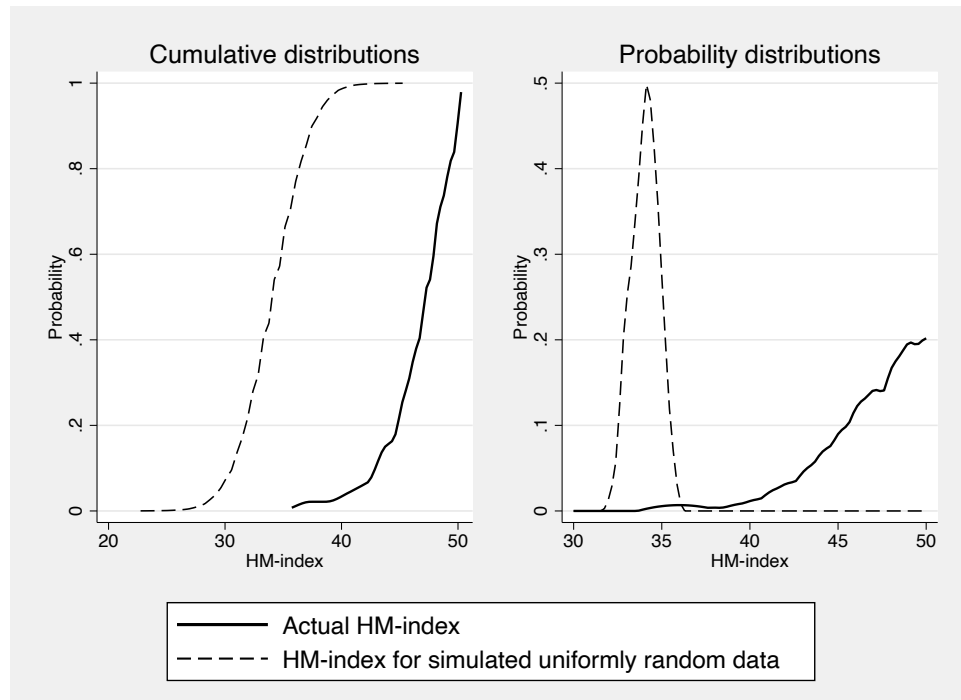9. Implemented using the *Stata* command `ksmirnov`.

Figure 2: *Left:* Kernel cumulative distribution functions of HM-indices from actual choices (solid line) and from simulated uniformly random data (dashed line) over all subjects. The kernel cdfs were generated using the `kdensity` package in Stata. *Right:* Kernel probability densities of HM-indices from actual choices (solid line) and from simulated uniformly random data (dashed line) over all subjects. The kernel pdfs were generated using the `kdens` package in Stata with the reflection option (since the HM-index is a bounded variable between 1 and 50)

Alston, J. M., and J. A. Chalfant. 1991. Can we take the con out of meat demand studies? *Western Journal of Agricultural Economics* 16: 36–48.

Banerjee, S., and J. H. Murphy. 2006. A simplified test for preference rationality of two-commodity choice. *Experimental Economics* 9: 67–75.

Choi, S., R. Fisman, D. Gale, and S. Kariv. 2007. Consistency and heterogeneity of individual behavior under uncertainty. *American Economic Review* 97: 1921–1938.

Demetry, M., P. Hjertstrand, and M. Polisson. 2022. Testing axioms of revealed preference in Stata. *Stata Journal* 22: ??–??

Gross, J., and D. Kaiser. 1996. Two simple algorithms for generating a subset of data consistent with WARP and other binary relations. *Journal of Business and Economic Statistics* 14: 251–255.

10

Heufer, J., and P. Hjertstrand. 2015. Consistent subsets: Computationally feasible methods to compute the Houtman–Maks-index. *Economics Letters* 128: 87–89.

Houtman, M., and J. Maks. 1985. Determining all maximal data subsets consistent with revealed preference. *Kwantitatieve Methoden* 19: 89–104.

Matzkin, R. L., and M. K. Richter. 1991. Testing strictly concave rationality. *Journal of Economic Theory* 53: 287–303.

Rose, H. 1958. Consistency of preference: The two-commodity case. *Review of Economic Studies* 25: 124–125.

Varian, H. R. 1982. The nonparametric approach to demand analysis. *Econometrica* 50: 945–974.

About the authors

Marcos Demetry is a PhD student at Linnaeus University in Växjö, Sweden, and an affiliated doctoral student at the Research Institute of Industrial Economics (IFN) in Stockholm, Sweden.

Per Hjertstrand is an Associate Professor and Research Fellow at the Research Institute of Industrial Economics (IFN) in Stockholm, Sweden.