FOR ONLINE PUBLICATION:
APPENDICES A-D FOR MEANINGFUL THEOREMS:
NONPARAMETRIC ANALYSIS OF REFERENCE-DEPENDENT PREFERENCES

Laura Blow
Department of Economics, University of Surrey

Vincent P. Crawford
Department of Economics, University of Oxford; All Souls College, Oxford;
and Department of Economics, University of California, San Diego

This version 24 November 2023

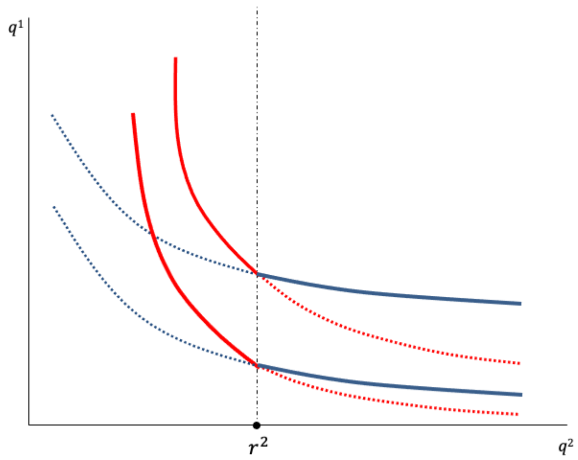
ONLINE APPENDIX A. Generalizing Tversky and Kahneman's notion of loss aversion with constant sensitivity and simplifying Proposition 4's sufficient conditions for a rationalization.

There is strong experimental and empirical support for loss aversion, whereby reference-dependent preferences are more sensitive to changes below a reference point than to equal changes above it (Kahneman and Tversky 1979, "KT"; Tversky and Kahneman 1991). We give a nonparametric generalization of Tversky and Kahneman's (1991, pp. 1047-1048) definition of loss aversion for the two-good case to the multi-good case. (KT considered only the one-good case, which is of limited interest in consumer theory.) Like Tversky and Kahneman we assume constant sensitivity, but we relax their assumption of additive separability across goods. The idea of loss aversion is still well defined with variable sensitivity, but formalizing it then is more complex, and Propositions 1 and 2 show that it would then be nonparametrically irrefutable anyway.

DEFINITION A1: [Preferences with constant sensitivity and loss aversion.] Assume that reference-dependent preferences and an associated utility function $u(\boldsymbol{q}, \boldsymbol{r})$ *have constant sensitivity. A collection of gain-loss regime preferences over consumption bundles satisfies loss aversion if and only if, for any observation* $\{\boldsymbol{p}_t, \boldsymbol{q}_t, \boldsymbol{r}_t\}$, *given* $\boldsymbol{r}_t$, *the preference ordering's global better-than-*$\boldsymbol{q}_t$ *set is weakly contained in each regime preference ordering's local better-than-*$\boldsymbol{q}_t$ *set.*

Figure A1 illustrates loss aversion with one active reference point and two gain-loss regimes. Loss aversion is a property of the relationship between regimes' preferences over consumption bundles given a reference point, thus independent of reference points themselves.

Figure A1. Loss aversion with one active reference point
(solid curves for active parts of indifference maps, dashed for inactive parts;
blue for gain map, red for loss map)



Because Definition A1's nesting of local and global better-than sets holds globally, loss aversion is equivalent to requiring the gain-loss regimes' indifference maps to satisfy a global single-crossing property: For any observation, across regimes that differ only in the gain-loss status of good *i*, the loss-side marginal rate of substitution between good *i* and any other good (generalized as needed for non-differentiable preferences) must be weakly more favorable to good *i* than the gain-side marginal rate of substitution. (Neoclassical preferences are thus weakly loss averse.) It is this single-crossing property, not the kinks in global indifference maps that it creates, that shapes loss aversion's nonparametric implications, which are testable with finite data. Loss aversion precludes nonconvex kinks, so if the regime maps all have convex better-than sets, then so do the associated global maps.

Corollary A1 shows that GARP for each regime's observations plus a condition weaker than loss aversion are sufficient for a rationalization. Recall that the gain-loss indicators $G_+^k(\boldsymbol{q},\boldsymbol{r}) = 1$ if $q_t^k \geq r_t^k$ and 0 otherwise and $G_-^k(\boldsymbol{q},\boldsymbol{r}) = 1$ if $q_t^k < r_t^k$ and 0 otherwise; and that $\Theta(\{\boldsymbol{q}_t,\boldsymbol{r}_t\}_{t=1,\dots,T}; g) \equiv \{t \in \{1,\dots,T\} \,|\, \boldsymbol{q}_t \in \Gamma(g;\boldsymbol{r}_t)\}$ is the set of observations t with $\boldsymbol{q}_t$ in regime g for $\boldsymbol{r}_t$.

COROLLARY A1: [Rationalization with modelable reference points via preferences and utility functions *with constant sensitivity that satisfy a condition weaker than loss aversion.] Suppose that reference-dependent preferences and an associated utility function are defined over K ≥ 2 goods, that reference-dependence is active for all K goods, that the preferences satisfy constant sensitivity and are continuous, and that the utility function therefore satisfies Proposition 3's (6). Consider data* $\{\boldsymbol{p}_t,\boldsymbol{q}_t,\boldsymbol{r}_t\}_{t=1,\dots,T}$ *with modelable reference points. If each gain-loss regime's data satisfy GARP within the regime; and there is some combination of preferences over consumption bundles, with continuous, strictly increasing consumption utility function* $U(\cdot)$ *and gain-loss component utility functions* $v_+^k(\cdot)$ *and* $v_-^k(\cdot)$*, such that, for any regime* g *and any pair of observations* $\sigma,\tau \in \Theta(\{\boldsymbol{q}_t,\boldsymbol{r}_t\}_{t=1,\dots,T}; g)$ *for which* $\boldsymbol{p}_\tau \cdot \boldsymbol{q}_\sigma \leq \boldsymbol{p}_\tau \cdot \boldsymbol{q}_\tau$ *(with the indicator functions* $G_+^k(\cdot,\cdot)$ *and* $G_-^k(\cdot,\cdot)$ *doing the work of a regime indicator function* $I_{\boldsymbol{g}}(\cdot,\cdot)$*),*

(A.1) $$U(\boldsymbol{q}_\sigma) + \textstyle\sum_k [G_+^k(\boldsymbol{q}_\sigma,\boldsymbol{r}_\tau) v_+^k(q_\sigma^k) + G_-^k(\boldsymbol{q}_\sigma,\boldsymbol{r}_\tau) v_-^k(q_\sigma^k)]$$
$$\leq \mathrm{U}(\boldsymbol{q}_\tau) + \textstyle\sum_k [G_+^k(\boldsymbol{q}_\tau,\boldsymbol{r}_\tau) v_+^k(q_\tau^k) + G_-^k(\boldsymbol{q}_\tau,\boldsymbol{r}_\tau) v_-^k(q_\tau^k)] + \lambda_\tau \boldsymbol{p}_\tau \cdot (\boldsymbol{q}_\sigma - \boldsymbol{q}_\tau),$$

and there are no observations for which $\boldsymbol{q}_t$ *is* not *on the boundary of the convex hull of* $\boldsymbol{q}_t$*'s upper contour set for the associated candidate global preference ordering for* $\boldsymbol{r}_t$*, then the consumption utility function* $U(\cdot)$ *and gain-loss component utility functions* $v_+^k(\cdot)$ *and* $v_-^k(\cdot)$ *rationalize the data.*

Proof: As in Proposition 4, by Afriat's Theorem, the hypothesized combination of preferences over bundles with consumption utility function $U(\cdot)$ and gain-loss component utility functions $v_+^k(\cdot)$ and $v_-^k(\cdot)$ prevent defections from any observation's consumption bundle to any affordable bundle in the same own gain-loss regime. If the hypothesized preferences are such that there are no observations t for which $\boldsymbol{q}_t$ is *not* on the boundary of the convex hull of the better-than-$\boldsymbol{q}_t$ set for the candidate global preference ordering given $\boldsymbol{r}_t$, then we can assume that they satisfy loss aversion without loss of generality. For, the candidate global ordering can then be replaced by a convexified ordering whose better-than-$\boldsymbol{q}_t$ sets are the convex hulls of the candidate global ordering, without changing any observation's optimal bundle. Definition A1 then implies that $U(\cdot)$ and the $v_+^k(\cdot)$ and $v_-^k(\cdot)$ also prevent defections from any observation's bundle to any affordable bundle in a different regime. Alternatively, consider a defection from $\boldsymbol{q}_\tau \in \Gamma(g; \boldsymbol{r}_\tau)$ to some $\boldsymbol{q} \in \Gamma(g'; \boldsymbol{r}_\tau)$ with $g' \neq g$ and $\boldsymbol{p}_\tau \cdot \boldsymbol{q} \leq \boldsymbol{p}_\tau \cdot \boldsymbol{q}_\tau$. If $\boldsymbol{q}$ were in regime g, we would have, by Afriat's Theorem,

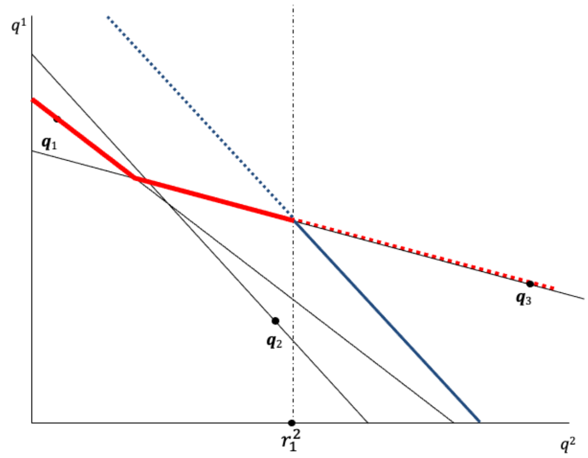
$$\text{(A.2)} \quad U(\boldsymbol{q}) + \textstyle\sum_k [G_+^k(\boldsymbol{q}, \boldsymbol{r}_\tau)\{v_+^k(q^k) - v_+^k(r_\tau^k)\} + G_-^k(\boldsymbol{q}, \boldsymbol{r}_\tau)\{v_-^k(q^k) - v_-^k(r_\tau^k)\}]$$
$$\leq U(\boldsymbol{q}_\tau) + \textstyle\sum_k [G_+^k(\boldsymbol{q}_\tau, \boldsymbol{r}_\tau)\{v_+^k(q_\tau^k) - v_+^k(r_\tau^k)\} + G_-^k(\boldsymbol{q}_\tau, \boldsymbol{r}_\tau)\{v_-^k(q_\tau^k) - v_-^k(r_\tau^k)\}].$$

Given that $\boldsymbol{q}$ is actually in regime g', the interpretation of loss aversion in terms of marginal rates of substitution implies that the left-hand side of (A.2) is lower or at least no higher than if $\boldsymbol{q}$ were in regime g. (A.2) thus prevents defections from $\boldsymbol{q}_\tau$ to affordable bundles in different regimes. ■

Corollary A1's final "no observations for which $\boldsymbol{q}_t$ is *not* on the boundary" condition rules out bunching of consumption bundles in regions of commodity space where the rationalizing regime preferences violate loss aversion, and is vacuously satisfied for preferences that satisfy loss aversion. Such restrictions on bunching are unusual in a nonparametric analysis. And although loss aversion is usually viewed as an empirically well-supported assumption with important implications, to our knowledge it has not previously been linked to the *existence* of a reference-dependent rationalization.

In Figure A2 the entire dataset violates GARP, the Afriat gain-loss regime preferences violate loss aversion, but the data satisfy Corollary A1's final conditions, thus allowing a rationalization. Only reference point $\boldsymbol{r}_1$ is shown and observation 1 is in the good-2 loss regime. Assume that $\boldsymbol{r}_2 = [0, 0]$, so that observation 2's budget set is entirely in the good-2 gain regime; and that $\boldsymbol{r}_3 = [0, m]$, where m is large enough that observation 3's budget set is entirely in the good-2 loss regime. The Afriat regime preferences yield a candidate for global preferences that make all three observations' consumption bundles optimal: Observations 2's and 3's budget sets are entirely in their regimes (good-2 gain and good-2 loss, respectively), so their bundles' optimality in their regimes suffices for global optimality. Observation 1's bundle is optimal for its good-2 loss regime preferences and Corollary 1 ensures that its bundle's optimality extends to its entire budget set.

Figure A2. Rationalizing data that violate GARP when preferences violate loss aversion but satisfy Corollary A1's sufficient conditions for a reationalization (solid curves for active parts of indifference maps, dashed for inactive parts; blue for gain map, red for loss map)



ONLINE APPENDIX B. Farber's (2005, 2008) dataset

Figure B.1: Hours and earnings choices, driver by driver

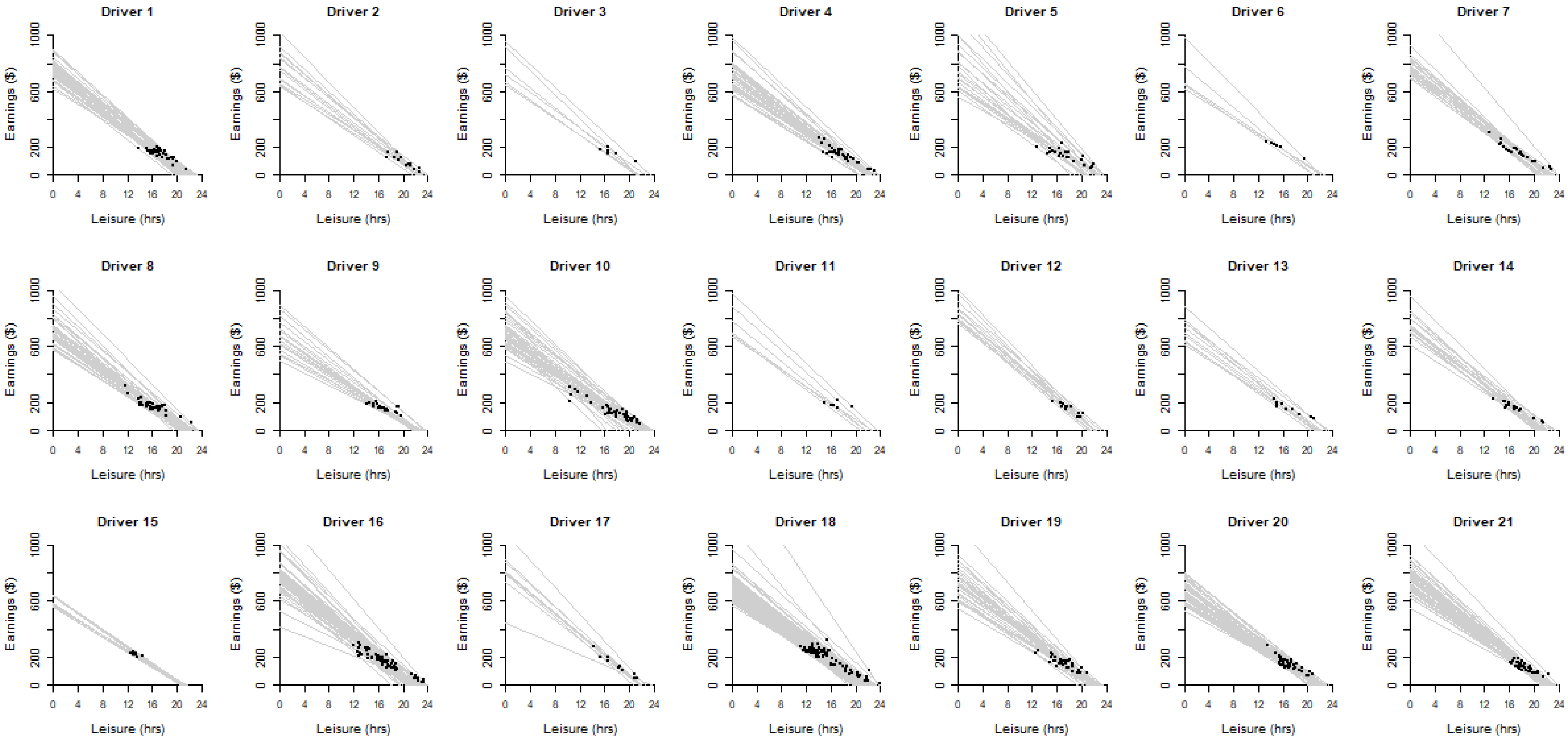


Table B.1: Descriptive statistics, driver by driver

	T	Working Hours	Driving Hours	Waiting Hours	Break Hours	Earnings ($/CPI)	Wage ($/hr)	Afriat Efficiency
Driver 1	39	6.85	4.32	2.53	0.90	153.01	36.41	0.9952
Driver 2	14	3.89	2.78	1.11	2.41	95.98	34.68	1
Driver 3	6	6.66	4.61	2.05	0.74	160.07	36.19	1
Driver 4	40	6.28	4.52	1.76	0.39	145.89	33.02	0.9978
Driver 5	23	6.46	3.98	2.48	2.11	144.00	38.12	0.9971
Driver 6	6	8.62	6.48	2.14	2.42	202.71	33.49	1
Driver 7	24	6.47	4.42	2.05	0.74	159.50	36.69	0.9991
Driver 8	37	7.78	5.13	2.64	0.86	170.33	34.23	0.9897
Driver 9	19	7.17	5.47	1.70	0.54	158.82	30.61	1
Driver 10	45	6.35	3.90	2.45	1.65	129.68	33.83	0.9954
Driver 11	6	7.15	5.22	1.93	0.71	182.40	35.50	1
Driver 12	13	6.15	4.03	2.13	0.55	155.57	39.44	0.9972
Driver 13	10	7.03	4.72	2.31	0.53	153.99	33.26	1
Driver 14	17	7.06	4.49	2.57	0.64	157.15	37.37	0.9930
Driver 15	8	10.82	7.64	3.17	0.19	217.29	29.92	1
Driver 16	70	6.84	4.56	2.28	0.93	163.56	37.72	0.9936
Driver 17	10	5.88	3.71	2.17	0.54	137.28	39.10	0.9946
Driver 18	72	8.53	5.84	2.69	0.60	194.88	35.07	0.9849
Driver 19	33	6.91	4.63	2.29	0.97	155.65	36.01	0.9870
Driver 20	46	7.10	4.80	2.30	0.67	148.76	32.73	0.9842
Driver 21	46	5.32	3.66	1.66	0.24	123.57	35.62	0.9915

ONLINE APPENDIX C. Pass Rates and Selten Measures for Reference-Dependent Models that Relax Additive Separability Across Goods

Figures C.1 and C.2 compare the empirical CDFs for different kinds of reference-point model (expectations- or experience-based, with various conditionings).

Figure C.1: Empirical CDFs of Proximities
for Different Kinds of Reference-dependent Model

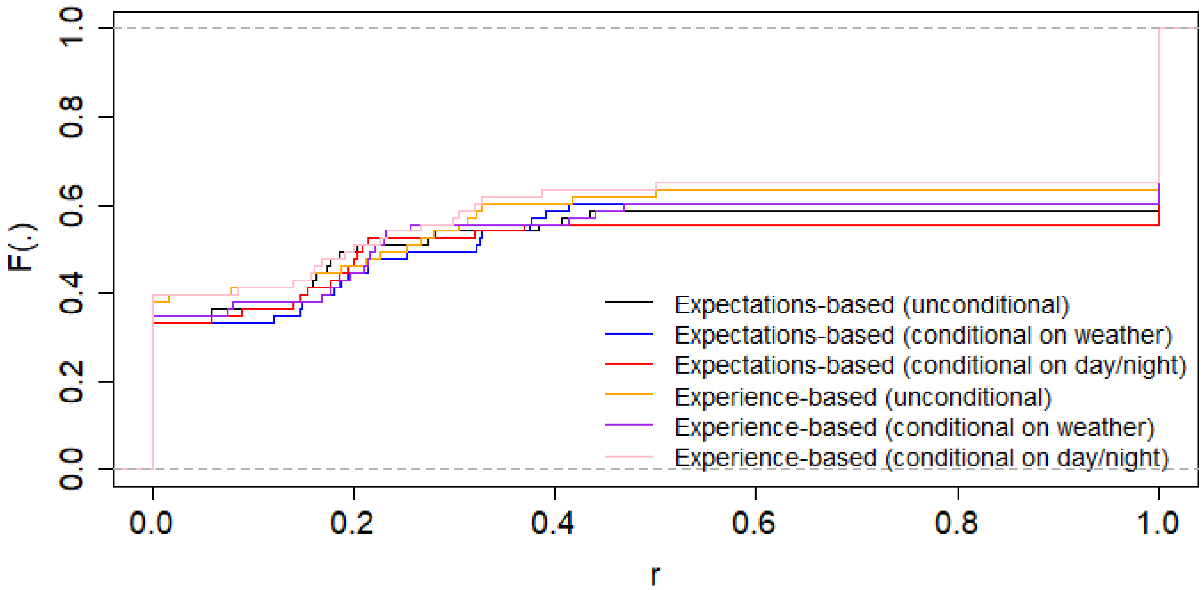
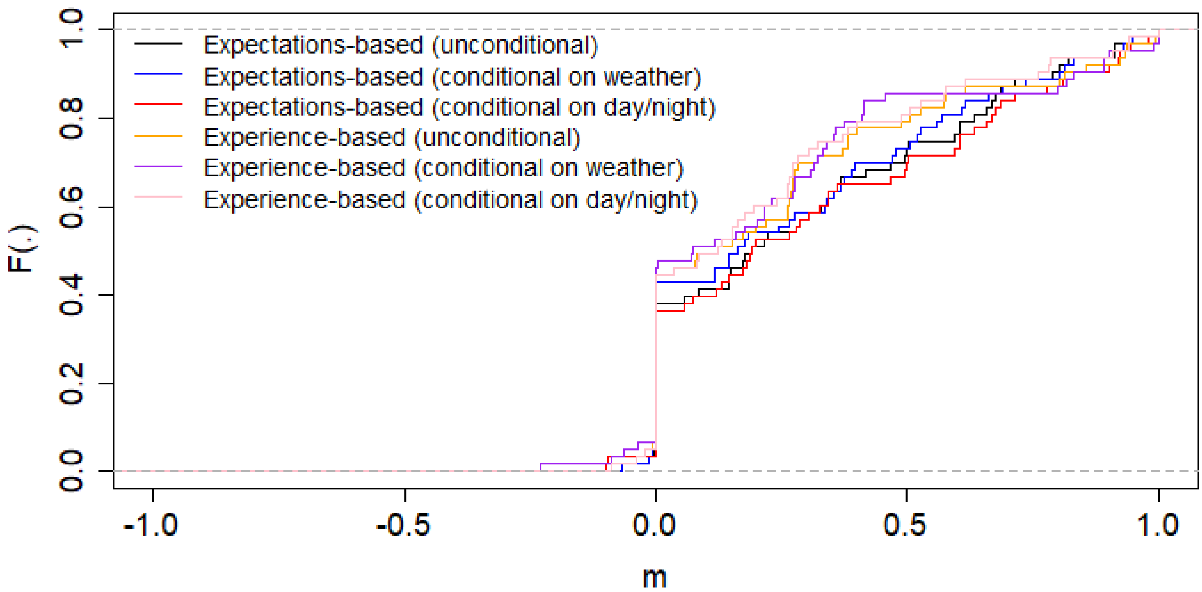


Figure C.2: Empirical CDFs of Selten Measures
for Different Kinds of Reference-dependent Model



Figures C.3 and C.4 compare the empirical CDFs for different forms of reference-dependence (hours only, earnings only, or both hours and earnings).

Figure C.3: Empirical CDFs of Proximities
for Different Forms of Reference-dependence

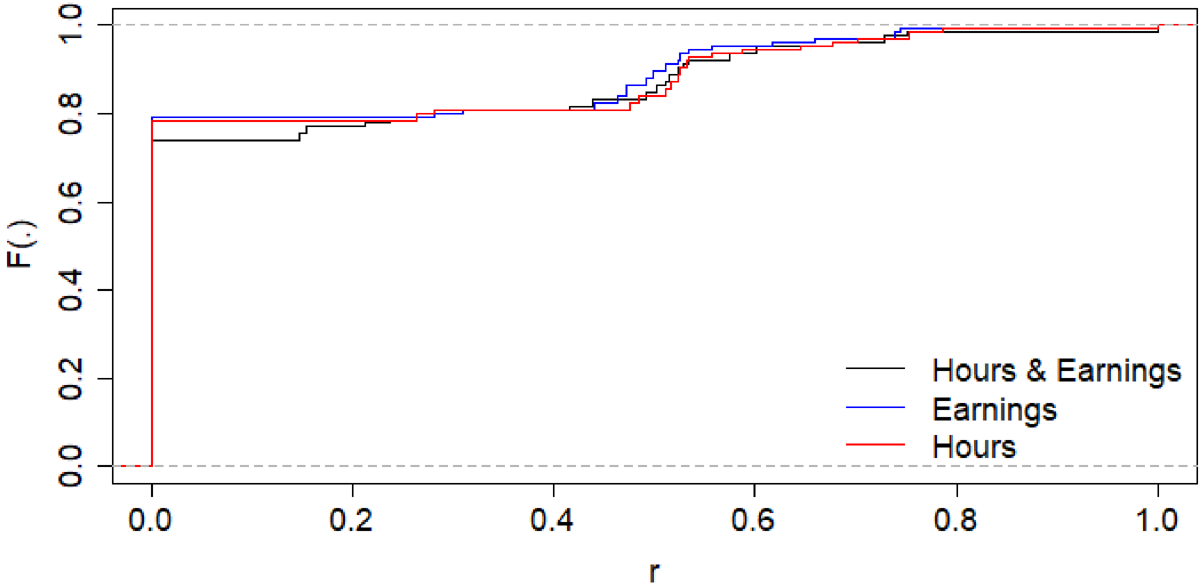
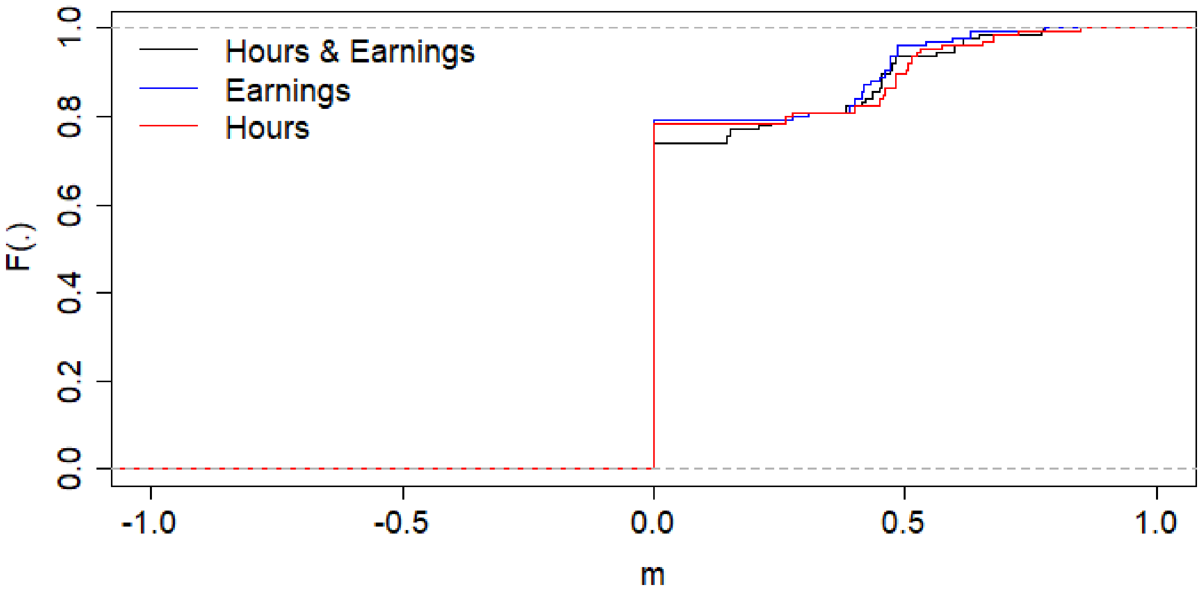


Figure C.4: Empirical CDFs of Selten Measures
for Different Forms of Reference-dependence



ONLINE APPENDIX D. Proximities and Selten Measures for Neoclassical and Unconditional Expectations- or Experience-based Reference-dependent Models that Relax Additive Separability Across Goods

Table D.1: Proximities for neoclassical and unconditional expectations-based reference-dependent models

	T	Neoclassical	Hours and Earnings	Earnings alone	Hours alone
Driver 1	39	0	0	0	1
Driver 2	14	1	1	1	1
Driver 3	6	1	1	1	1
Driver 4	40	0	0	0	0
Driver 5	23	0.048492	0.384315	0.406994	0.43469
Driver 6	6	1	1	1	1
Driver 7	24	0	0.162525	0.203046	0.15848
Driver 8	37	0	0.146842	0	0.17589
Driver 9	19	1	1	1	1
Driver 10	45	0	0	0	0
Driver 11	6	1	1	1	1
Driver 12	13	0.256557	1	0.15315	0.13958
Driver 13	10	1	1	1	1
Driver 14	17	0.265625	0.172794	0.280798	0.27437
Driver 15	8	1	1	1	1
Driver 16	70	0	0	0	0
Driver 17	10	0.207382	1	1	1
Driver 18	72	0	0	0	0
Driver 19	33	0	0.185454	0.057727	0
Driver 20	46	0	0	0	0
Driver 21	46	0	0	0	0

Table D.2: Selten measures for neoclassical and unconditional expectations-based reference-dependent models[1]

	T	Neoclassical	Hours and Earnings	Earnings alone	Hours alone
Driver 1	39	0	0	0	1
Driver 2	14	0.942	0.788	0.803	0.82
Driver 3	6	0.505	0.343	0.344	0.361
Driver 4	40	0	0	0	0
Driver 5	23	0.047492	0.328315	0.367994	0.416699
Driver 6	6	0.665	0.466	0.499	0.497
Driver 7	24	0	0.148525	0.193046	0.145481
Driver 8	37	0	0.144842	0	0.174891
Driver 9	19	0.987	0.912	0.913	0.941
Driver 10	45	0	0	0	0
Driver 11	6	0.421	0.303	0.287	0.279
Driver 12	13	0.187557	0.714	-0.09685	-0.09742
Driver 13	10	0.832	0.659	0.632	0.669
Driver 14	17	0.261625	0.084794	0.217798	0.222375
Driver 15	8	0.8	0.504	0.604	0.594
Driver 16	70	0	0	0	0
Driver 17	10	-0.01062	0.677	0.604	0.685
Driver 18	72	0	0	0	0
Driver 19	33	0	0.178454	0.056727	-0.002
Driver 20	46	0	0	0	0
Driver 21	46	0	0	0	0

[1] The Selten measures are Section III.D's lower bounds, the estimates imposing Proposition 5's full conditions [A]-[C].

Table D.3: Proximities for neoclassical and unconditional experience-based reference-dependent models

	T	Neoclassical	Hours and Earnings	Earnings alone	Hours alone
Driver 1	39	0	0	0	0
Driver 2	14	1	1	1	1
Driver 3	6	1	1	1	1
Driver 4	40	0	0.078039	0	0
Driver 5	23	0.048492	0.016492	0.212637	0.187298
Driver 6	6	1	1	1	1
Driver 7	24	0	1	0.278643	0.157611
Driver 8	37	0	1	0	0
Driver 9	19	1	1	1	1
Driver 10	45	0	0	0	0
Driver 11	6	1	1	1	1
Driver 12	13	0.256557	0.13873	0.265933	0.226169
Driver 13	10	1	1	1	1
Driver 14	17	0.265625	0.313394	0.416434	0.251828
Driver 15	8	1	1	1	1
Driver 16	70	0	0	0	0
Driver 17	10	0.207382	0.500883	0.326962	0.304039
Driver 18	72	0	0	0	0
Driver 19	33	0	0.322158	0	0
Driver 20	46	0	0	0	0
Driver 21	46	0	0	0	0

Table D.4: Selten measures for neoclassical and unconditional experience-based reference-dependent models[2]

	T	Neoclassical	Hours and Earnings	Earnings alone	Hours alone
Driver 1	39	0	0	0	0
Driver 2	14	0.942	0.804	0.856	0.814
Driver 3	6	0.505	0.373	0.401	0.381
Driver 4	40	0	0.077039	0	0
Driver 5	23	0.047492	-0.00751	0.197637	0.175298
Driver 6	6	0.665	0.267	0.274	0.285
Driver 7	24	0	0.994	0.276643	0.153611
Driver 8	37	0	1	0	0
Driver 9	19	0.987	0.924	0.935	0.938
Driver 10	45	0	0	0	0
Driver 11	6	0.421	0.271	0.261	0.261
Driver 12	13	0.187557	-0.08827	0.081933	0.036169
Driver 13	10	0.832	0.575	0.616	0.577
Driver 14	17	0.261625	0.270394	0.381434	0.219828
Driver 15	8	0.8	0.527	0.505	0.49
Driver 16	70	0	0	0	0
Driver 17	10	-0.01062	0.124883	-0.02104	-0.03996
Driver 18	72	0	0	0	0
Driver 19	33	0	0.321158	0	0
Driver 20	46	0	0	0	0
Driver 21	46	0	0	0	0

[2] The Selten measures are Section III.D's lower bounds, the estimates imposing Proposition 5's full conditions [A]-[C].