SUPPLEMENTARY MATERIAL FOR "SIMULTANEOUS MEAN-VARIANCE REGRESSION"

RICHARD H. SPADY[†] AND SAMI STOULI[§]

1. Summary

This Supplementary Material presents further simulation results and an additional empirical example for "Simultaneous Mean-Variance Regression".

In Section 2 we give additional results for the simulations based on MacKinnon (2013) in order to study further the finite-sample properties of MVR. We compare the finite-sample inference performance of MVR, OLS and WLS when finite-sample corrections are applied to the standard errors of each estimator. For completeness we also replicate all experiments under the assumption that the conditional mean function (CMF) is correctly specified and the variance is misspecified, imposing the simplifications shown in equation (4.3) in the calculation of the MVR standard errors in Theorem 6 of the main text. Overall we find that the favorable theoretical properties of MVR translate into very substantial finite-sample gains over both OLS and WLS in terms of estimation performance and largely improved inference. In summary, for the numerical simulations based on MacKinnon (2013) the main findings are:

- MVR-based inference brings very large improvements relative to inference based on the asymptotic heteroskedasticity-robust standard errors. We find that in the presence of heteroskedasticity rejection probabilities for MVR are much closer to nominal level than those for OLS and for WLS with misspecified weights. These relative gains of MVR remain large when finite-sample corrections of the standard errors are implemented for all estimators.
- MVR achieves the improvements above while simultaneously displaying tighter confidence intervals in all designs for sample sizes large enough. When MVR standard errors are calculated under the assumption that the CMF is correctly

Date: January 2, 2019.
[†] Nuffield College, Oxford, and Department of Economics, Johns Hopkins University, rspady@jhu.edu.
§ Department of Economics, University of Bristol, s.stouli@bristol.ac.uk.

specified, then average MVR confidence interval lengths across simulations are shorter than their OLS counterpart for all sample sizes and designs. They are also shorter than their WLS counterpart with misspecified weights for sample sizes large enough. The relative gains of MVR are then also larger when finite-sample corrections of the standard errors are implemented for all estimators.
- The precision of MVR estimates in root mean squared error (RMSE) is largely superior to OLS under heteroskedasticity and to WLS with misspecified weights, with lower losses than WLS relative to OLS under homoskedasticity.

In Section 3 we give a full set of additional results for the reversal of fortune application, reporting standard errors with finite-sample adjustments and with and without assuming correct specification of the CMF. We find that in this example MVR standard errors are robust to finite-sample corrections and standard errors assuming correct specification of the CMF tend to be slightly smaller.

In Section 4, we report an additional empirical application to demand for gasoline in the United States, and additional numerical simulations calibrated to this example. In particular, we study the finite-sample approximation properties of MVR by implementing simulations calibrated to the demand for gasoline empirical example with a nonlinear CMF. Overall, all experiments confirm the favorable finite-sample estimation, inference and approximation properties of MVR.

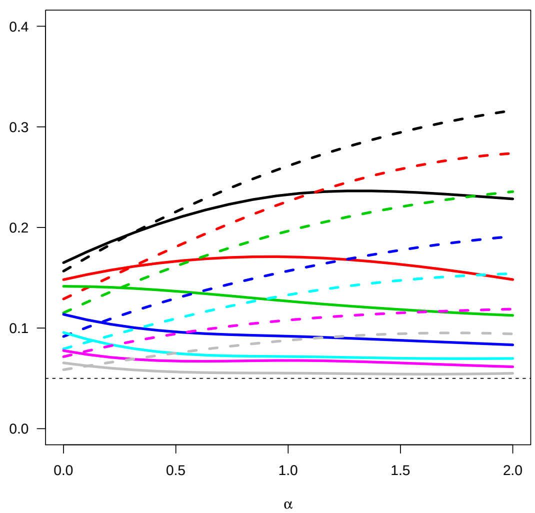
2. Additional Results for the Numerical Simulations Based on MacKinnon (2013)

2.1. **Finite-Sample Corrections.** In this Section we compare rejection probabilities and confidence intervals based on standard errors with finite-sample adjustments proposed by MacKinnon and White (1985), for the numerical simulations in Section 5 of the main text. For an estimator $\tilde{\beta}$ of β, define the sample residuals $\tilde{u}_i := y_i - x_i'\tilde{\beta}$, $i = 1, \dots, n$. The first correction (HC1) uses a degrees-of-freedom adjustment by rescaling the squared sample residuals by the factor $n/(n-k)$ in the calculation of the heteroskedasticity-robust standard errors. The second correction we consider approximates a jackknife estimator (HC3) for the robust variance-covariance matrix, as suggested for small samples by Long and Ervin (2000), for instance, where for each $i = 1, \dots, n$, the ith squared sample residual is rescaled by the factor $1/(1-h_i)$, where h_i is the ith element of the hat matrix $X_n(X_n'X_n)^{-1}X_n'$. The same corrections are applied to the standard errors of OLS, WLS and MVR.

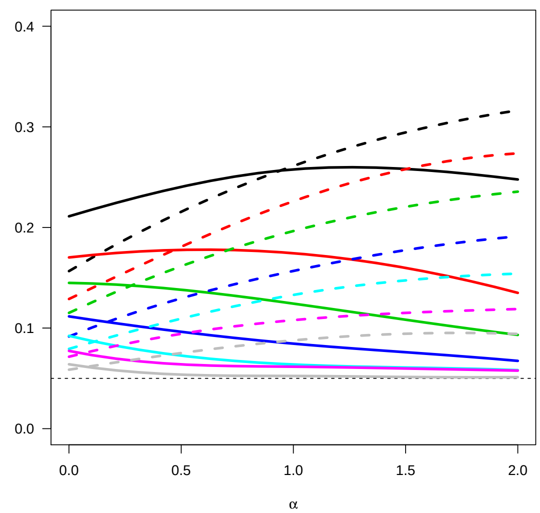
As in the main text we consider rejection probabilities for β_4. Figure 2.1 shows that when degrees-of-freedom corrections (HC1) are implemented, the OLS and WLS rejection probabilities improve but MVR-based inference continues to yield large improvements for all sample sizes in the presence of heteroskedasticity, with the exception of OLS rejection probability for $n = 20$ and $\alpha = 0.5$ which is slightly lower than its e-MVR counterpart (Figure 2.1(b)). Figures 2.2-2.3 compare rejection probability curves when the HC3 correction is implemented. For clarity of representation the curves for MVR and OLS/WLS are shown on different figures and the scale has been modified compared to Figure 2.1. Overall, the rejection probabilities are largely reduced for all estimators, and MVR-based inference continues to yield substantial improvements in the presence of heteroskedasticity relative to both OLS and WLS. The main exceptions are for $n = 20$, and partly for $n = 40$, where the OLS rejection probability curves are not placed above the other curves and show rejection probabilities closer to the nominal level than the corresponding MVR curves (Figures 2.2(b) and (d)).

In order to further investigate the relative performance of MVR-based inference with finite-sample corrections, Tables 1-4 report the ratio of average MVR confidence interval lengths across simulations over the average OLS and WLS confidence interval lengths for β_4 for each sample size and value of heteroskedasticity index α, in percentage terms. Tables 1 and 3 show that the relatively larger average length of the confidence intervals for ℓ-MVR when $n = 20$ in Tables 4-5 in the main text is very much reduced with finite-sample corrections. With HC3 corrections, Tables 2 and 4 show that in the presence of heteroskedasticity shorter average MVR confidence interval lengths across simulations are obtained for all sample sizes and all designs, except for e-MVR with $\alpha = 0.5$. These results show that finite-sample corrections substantially reduce the average MVR confidence interval lengths relative to both OLS and WLS.

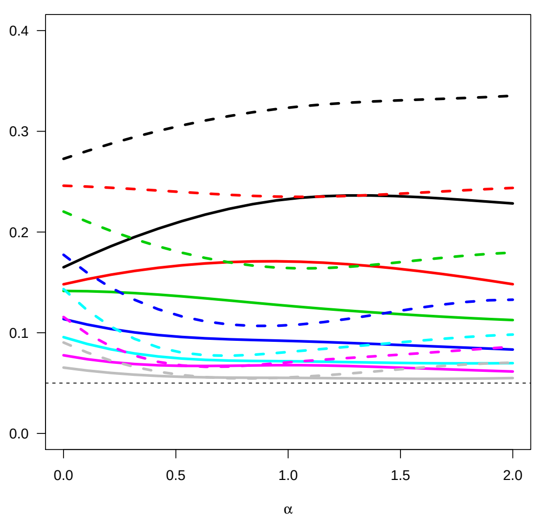
Although the finite-sample corrections for MVR should be regarded as experimental, the simulation results we report indicate that additional improvements relative to OLS and WLS can be achieved and should be explored further in future work.



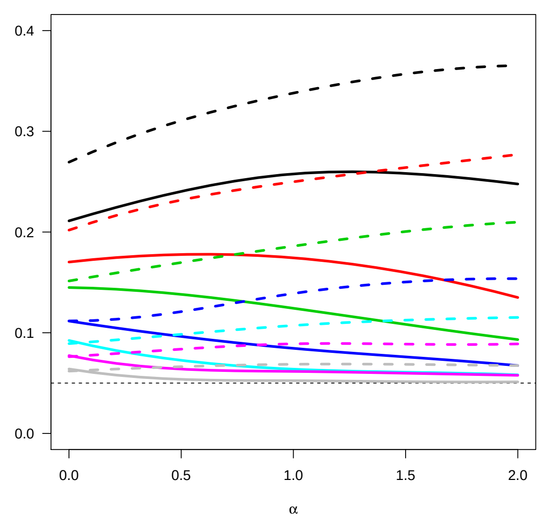
(A) ℓ-MVR vs OLS.



(B) e-MVR vs OLS.

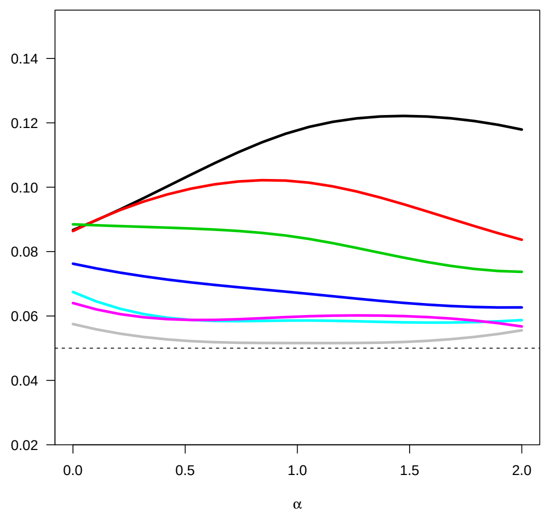


(C) ℓ-MVR vs ℓ-WLS.

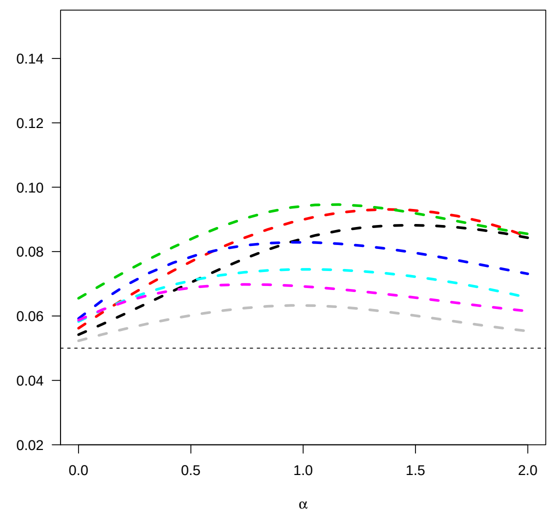


(D) e-MVR vs e-WLS.

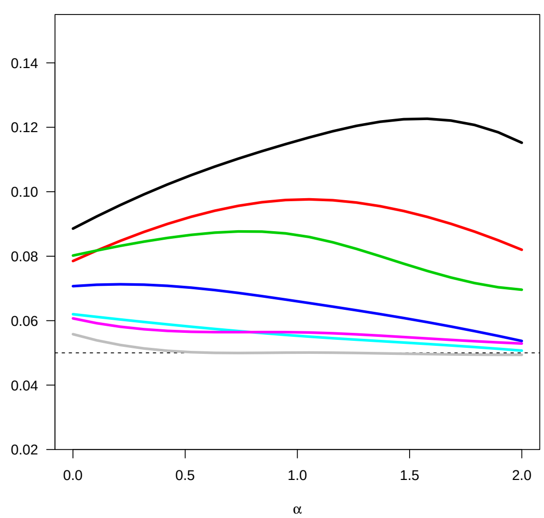
FIGURE 2.1. Rejection frequencies for asymptotic t tests calculated with standard errors with HC1 correction: MVR (solid lines), and OLS and WLS (dashed lines). Sample sizes: 20 (black), 40 (red), 80 (green), 160 (blue), 320 (cyan), 640 (magenta), 1280 (grey).



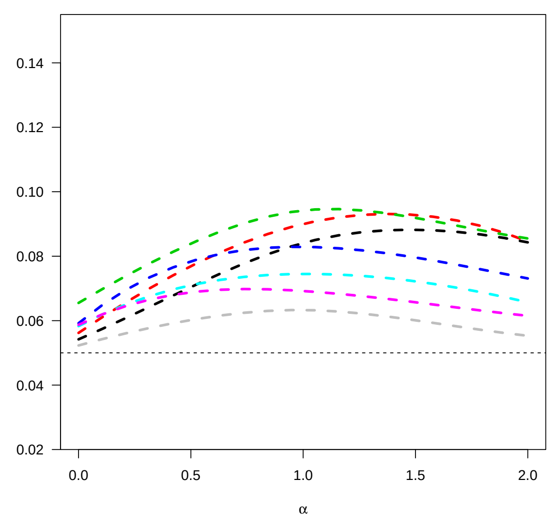
(A) ℓ-MVR.



(B) OLS.

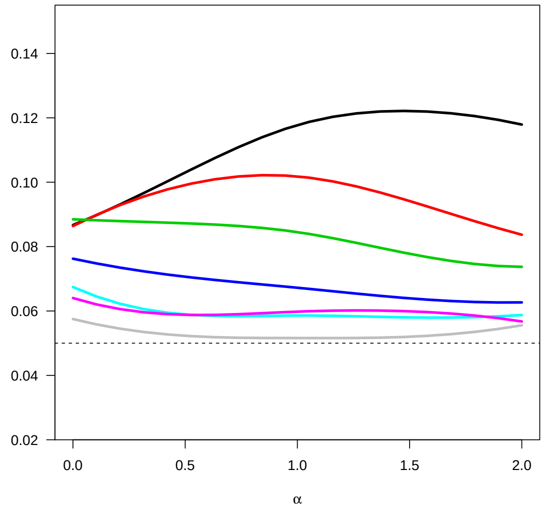


(C) e-MVR.

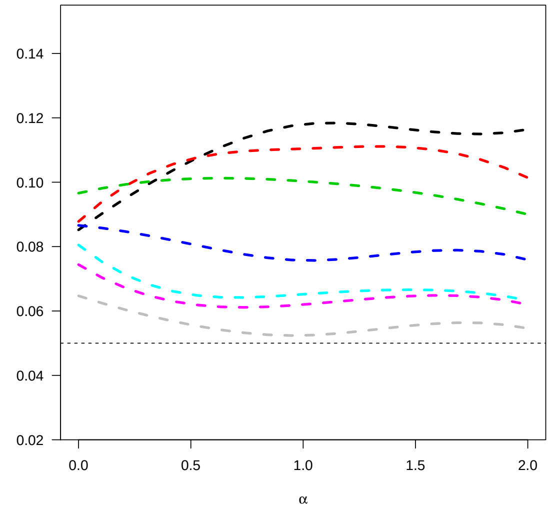


(D) OLS.

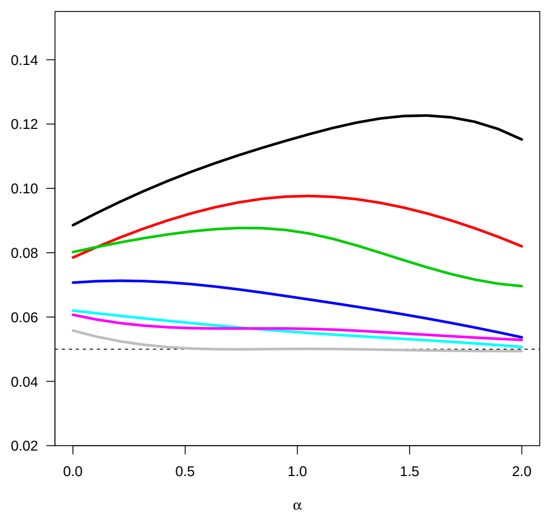
FIGURE 2.2. Rejection frequencies for asymptotic t tests calculated with standard errors with HC3 correction: MVR (solid lines), and OLS (dashed lines). Sample sizes: 20 (black), 40 (red), 80 (green), 160 (blue), 320 (cyan), 640 (magenta), 1280 (grey).



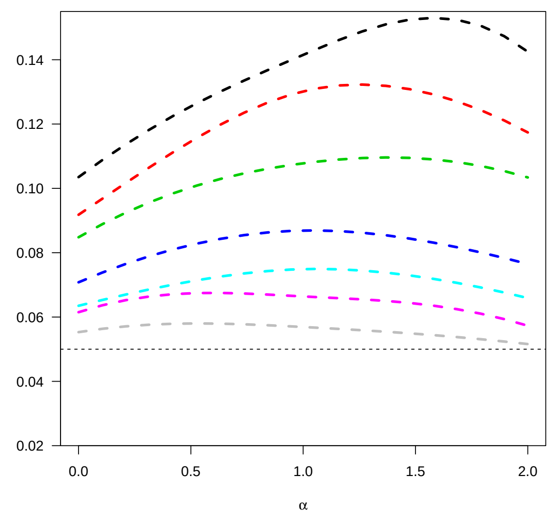
(A) ℓ-MVR.



(B) ℓ-WLS.



(C) e-MVR.



(D) e-WLS.

FIGURE 2.3. Rejection frequencies for asymptotic t tests calculated with standard errors with HC3 correction: MVR (solid lines), and WLS (dashed lines). Sample sizes: 20 (black), 40 (red), 80 (green), 160 (blue), 320 (cyan), 640 (magenta), 1280 (grey).

	ℓ-MVR					e-MVR				
α	0	0.5	1	1.5	2	0	0.5	1	1.5	2
$n = 20$	109.8	111.0	109.4	105.3	100.7	101.7	106.5	107.4	105.2	101.0
$n = 40$	106.0	107.7	103.6	95.6	87.0	101.7	107.0	105.9	99.9	90.8
$n = 80$	102.6	104.3	97.0	85.1	74.1	101.0	105.3	100.0	88.5	73.9
$n = 160$	101.4	101.5	90.0	74.8	63.3	100.7	103.0	92.7	76.3	59.1
$n = 320$	100.8	98.5	83.0	65.5	54.7	100.5	100.0	84.8	65.1	47.3
$n = 640$	100.4	95.2	76.2	57.4	47.4	100.2	96.3	77.0	55.6	38.1
$n = 1280$	100.3	92.1	70.1	50.5	41.1	100.2	93.0	70.3	47.9	31.1

TABLE 1. Ratio ($\times 100$) of MVR average confidence interval lengths for β_4 over corresponding OLS counterpart. Confidence intervals constructed with standard errors with HC1 correction.

	ℓ-MVR					e-MVR				
α	0	0.5	1	1.5	2	0	0.5	1	1.5	2
$n = 20$	93.5	90.5	84.8	77.9	71.3	93.4	91.1	86.0	79.0	71.5
$n = 40$	97.9	93.9	84.9	74.2	64.2	98.6	94.7	86.0	75.3	63.9
$n = 80$	99.8	94.2	82.2	68.3	56.9	100.0	94.5	82.8	68.4	54.2
$n = 160$	100.5	93.0	77.7	61.5	50.0	100.4	93.3	78.1	60.8	45.1
$n = 320$	100.5	91.3	73.0	55.1	44.3	100.5	91.7	73.2	53.6	37.5
$n = 640$	100.3	89.5	68.7	49.8	39.6	100.3	89.9	68.6	47.7	31.7
$n = 1280$	100.3	87.7	64.4	44.8	35.1	100.2	88.0	64.1	42.2	26.8

TABLE 2. Ratio ($\times 100$) of MVR average confidence interval lengths for β_4 over corresponding OLS counterpart. Confidence intervals constructed with standard errors with HC3 correction.

	ℓ-MVR					e-MVR				
α	0	0.5	1	1.5	2	0	0.5	1	1.5	2
$n = 20$	131.9	129.5	128.1	126.8	125.3	120.7	125.3	130.5	135.3	137.5
$n = 40$	123.1	118.1	116.3	114.9	111.9	110.3	115.8	122.6	128.6	129.2
$n = 80$	115.9	109.8	109.1	107.0	102.0	105.0	110.0	115.3	118.6	112.8
$n = 160$	111.0	105.4	106.2	101.8	95.7	102.6	106.5	108.4	107.0	96.0
$n = 320$	108.0	103.4	105.4	98.2	91.7	101.3	103.8	101.9	96.5	81.7
$n = 640$	105.3	102.7	105.1	95.3	88.2	100.6	101.1	95.7	87.0	70.1
$n = 1280$	103.4	102.7	104.9	93.0	85.0	100.4	98.9	90.5	79.4	60.9

Table 3. Ratio (×100) of MVR average confidence interval lengths for β_4 over corresponding WLS counterpart. Confidence intervals constructed with standard errors with HC1 correction.

	ℓ-MVR					e-MVR				
α	0	0.5	1	1.5	2	0	0.5	1	1.5	2
$n = 20$	100.7	102.5	103.0	101.6	98.3	105.8	108.6	111.5	112.2	109.0
$n = 40$	101.3	105.3	106.2	103.5	97.7	104.7	107.5	110.1	109.9	103.9
$n = 80$	101.8	105.9	106.7	101.7	93.5	103.0	104.2	105.2	103.0	93.1
$n = 160$	101.7	105.6	106.6	98.6	89.3	101.9	101.2	99.3	94.1	81.0
$n = 320$	101.2	104.9	106.1	96.0	86.6	101.1	99.2	94.6	86.6	70.7
$n = 640$	100.9	104.1	105.7	93.4	82.8	100.6	97.5	90.1	79.6	62.0
$n = 1280$	100.6	103.6	105.3	91.1	79.2	100.4	96.0	86.4	74.2	55.4

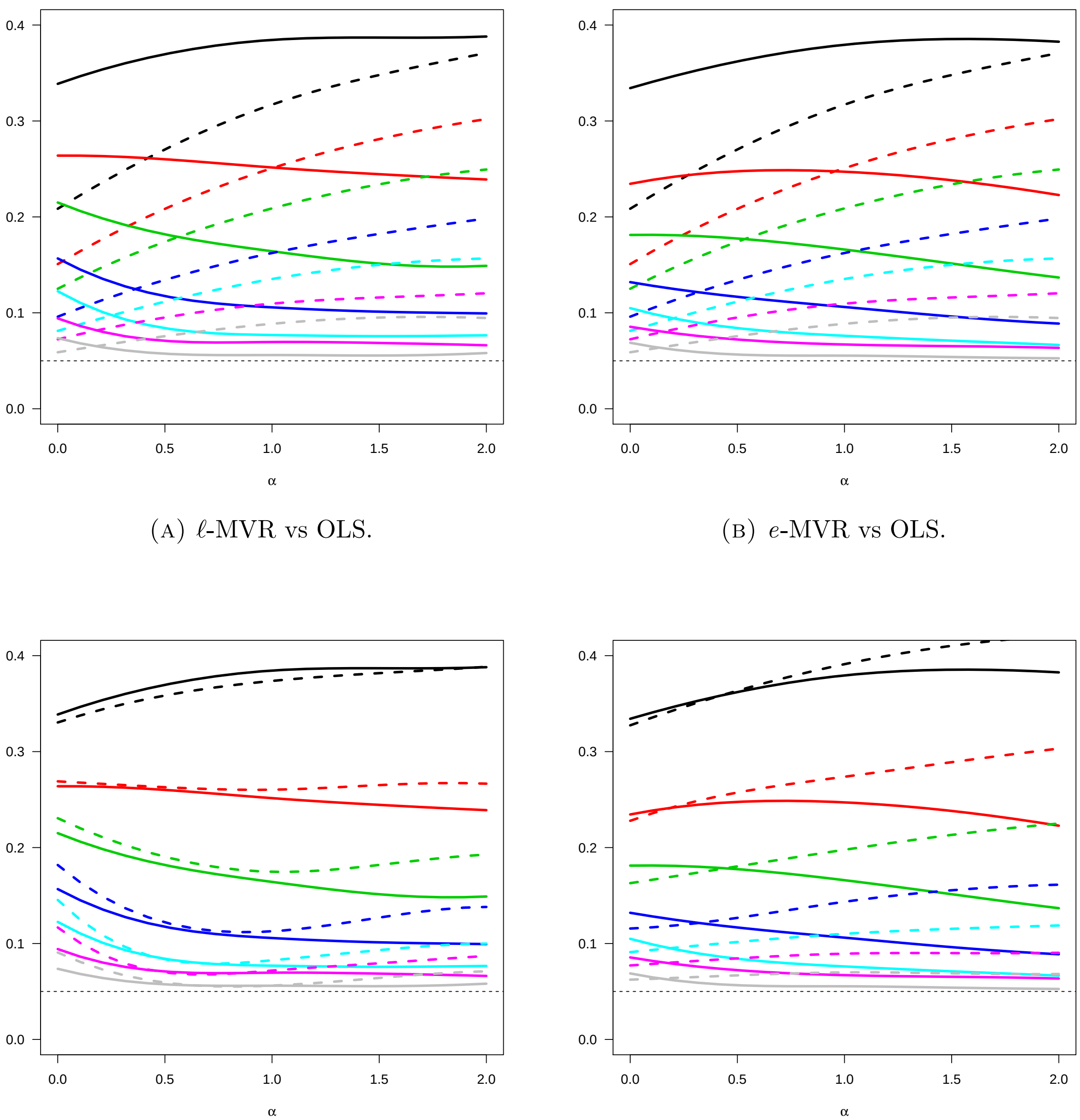
Table 4. Ratio (×100) of MVR average confidence interval lengths for β_4 over corresponding WLS counterpart. Confidence intervals constructed with standard errors with HC3 correction.

2.2. **Inference under Variance Misspecification.** In order to complete this study of the finite-sample performance of MVR inference relative to heteroskedasticity-robust OLS and WLS inference, we also compare the rejection probabilities and the lengths of the confidence intervals when MVR standard errors are calculated under the assumption of correct specification of the CMF, imposing the simplifications shown in equation (4.3) in the main text.

Figure 2.4 shows that the rejection probability curves for ℓ-MVR and $n = 20, 40$ are now placed well above the curves for larger sample sizes. MVR leads to smaller rejection probabilities than OLS for $\alpha \geq 1$ with $n = 40$, and in the presence of heteroskedasticity for all larger sample sizes. The MVR curves are now closer to WLS curves although the overall improvements remain substantial, in particular relative to e-WLS. The finite-sample corrections results in Figures 2.5-2.7 do not alter substantially the main conclusions.

In terms of relative confidence interval lengths, Tables 5-7 show that assuming correct specification of the CMF lead to average MVR confidence interval lengths that are shorter for all sample sizes and designs compared to OLS. The degrees-of-freedom corrections HC1 do not affect the relative length of the confidence intervals. Compared to WLS, Tables 8-10 again show shorter average MVR confidence interval lengths for n large enough and all designs where the conditional variance function is misspecified.

Overall these simulation results under correct specification of the CMF further illustrate the large MVR finite-sample improvements for inference in heteroskedastic designs.



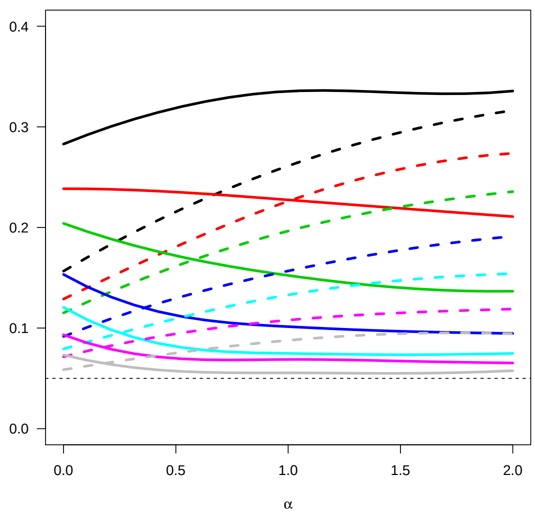
(A) ℓ-MVR vs OLS.

(B) e-MVR vs OLS.

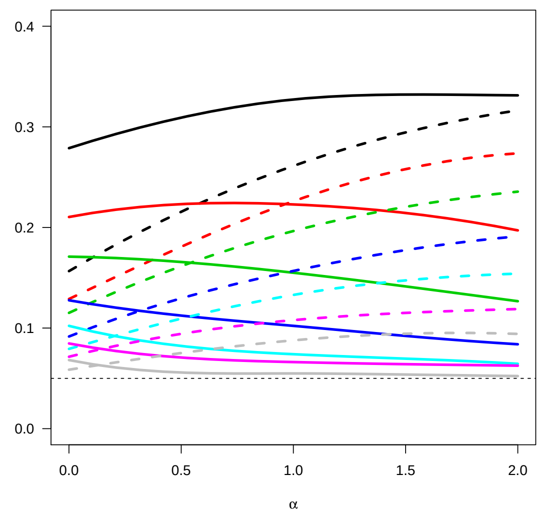
(C) ℓ-MVR vs ℓ-WLS.

(D) e-MVR vs e-WLS.

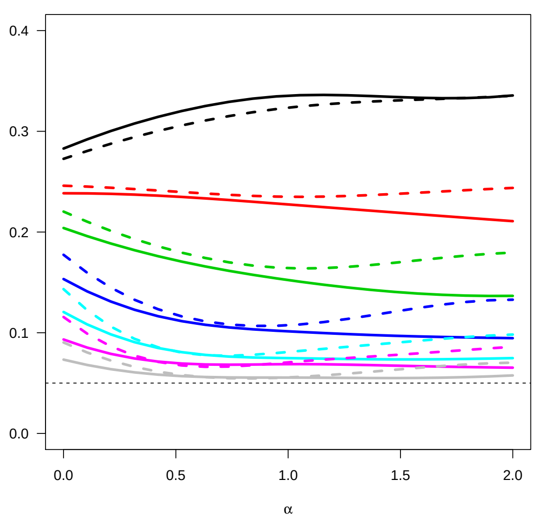
FIGURE 2.4. Rejection frequencies for asymptotic t tests calculated with asymptotic standard errors under correct specification of the CMF: MVR (solid lines), and OLS and WLS (dashed lines). Sample sizes: 20 (black), 40 (red), 80 (green), 160 (blue), 320 (cyan), 640 (magenta), 1280 (grey).



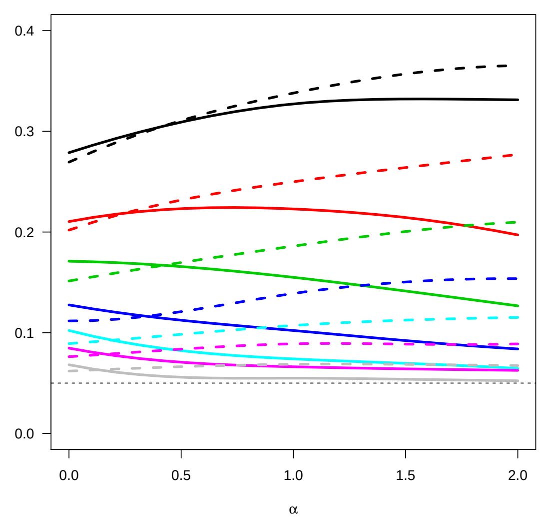
(A) ℓ-MVR vs OLS.



(B) e-MVR vs OLS.

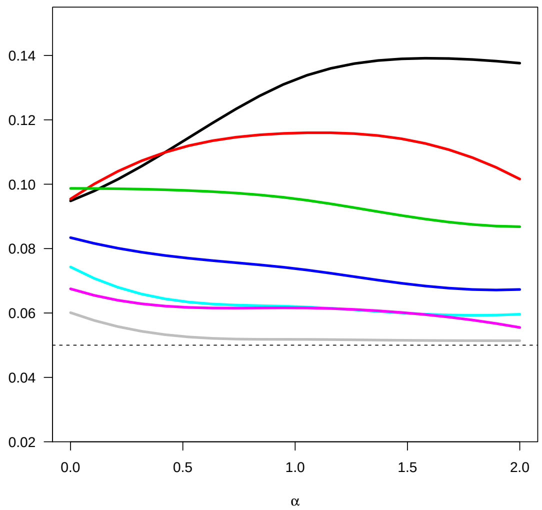


(C) ℓ-MVR vs ℓ-WLS.

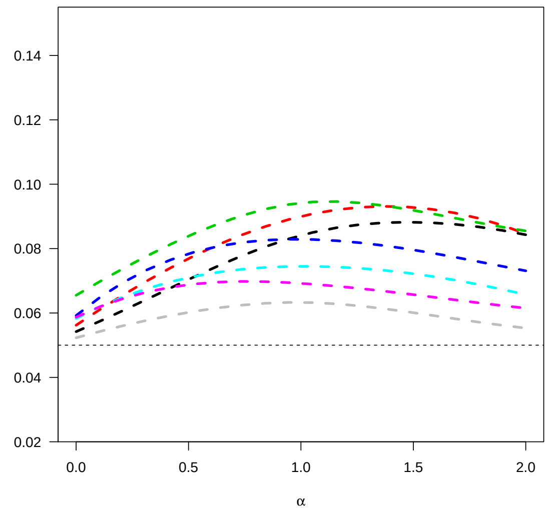


(D) e-MVR vs e-WLS.

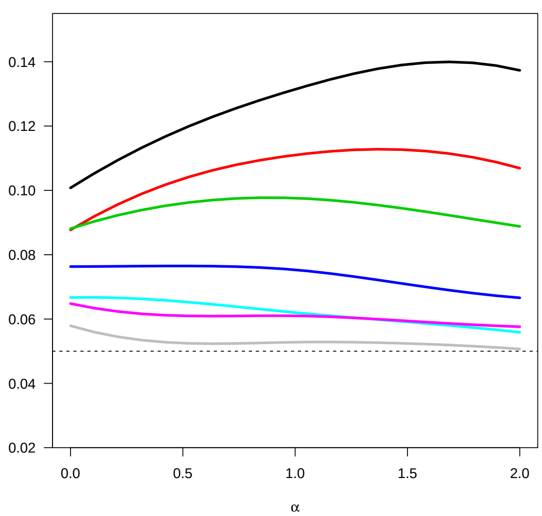
FIGURE 2.5. Rejection frequencies for asymptotic t tests calculated with standard errors with HC1 correction, under correct specification of the CMF: MVR (solid lines), and OLS and WLS (dashed lines). Sample sizes: 20 (black), 40 (red), 80 (green), 160 (blue), 320 (cyan), 640 (magenta), 1280 (grey).



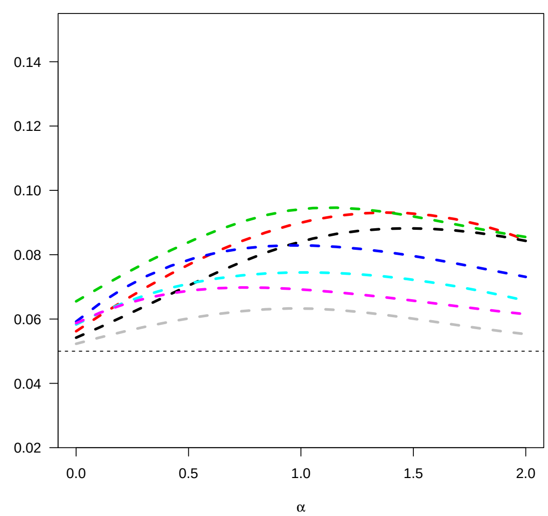
(A) ℓ-MVR.



(B) OLS.

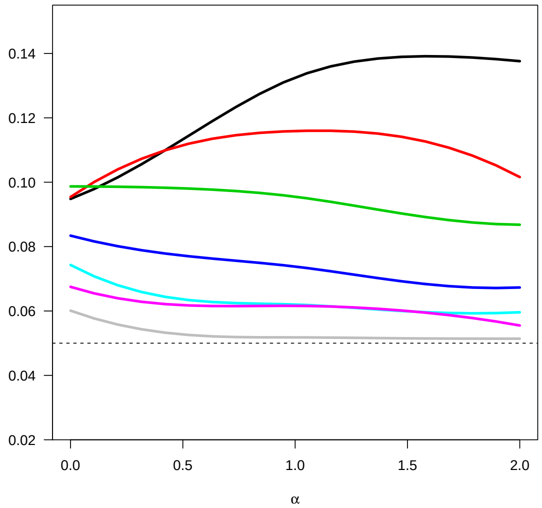


(C) e-MVR.

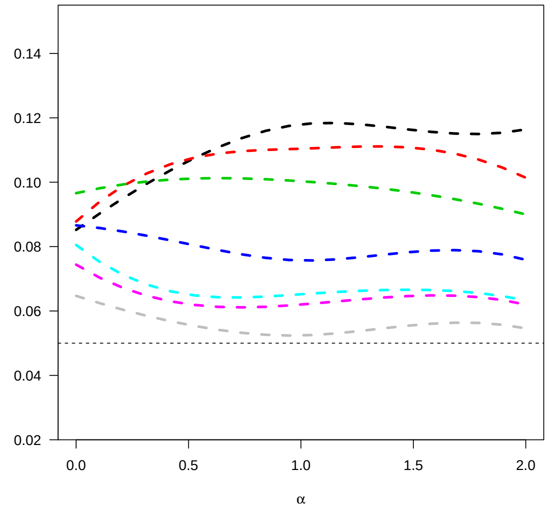


(D) OLS.

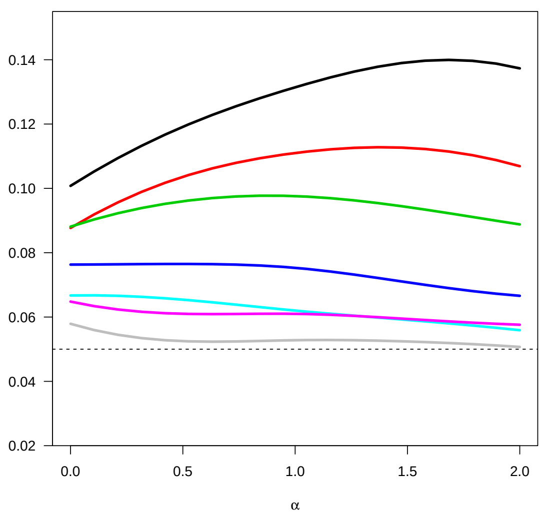
FIGURE 2.6. Rejection frequencies for asymptotic t tests calculated with standard errors with HC3 correction, under correct specification of the CMF: MVR (solid lines), and OLS and WLS (dashed lines). Sample sizes: 20 (black), 40 (red), 80 (green), 160 (blue), 320 (cyan), 640 (magenta), 1280 (grey).



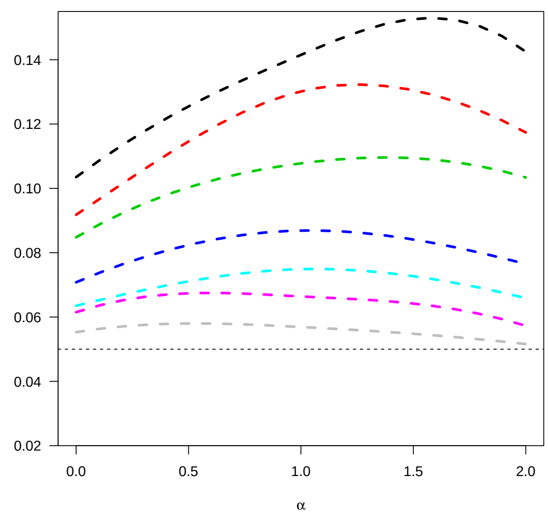
(A) ℓ-MVR.



(B) ℓ-WLS.



(C) e-MVR.



(D) e-WLS.

FIGURE 2.7. Rejection frequencies for asymptotic t tests calculated with standard errors with HC3 correction, under correct specification of the CMF: MVR (solid lines), and OLS and WLS (dashed lines). Sample sizes: 20 (black), 40 (red), 80 (green), 160 (blue), 320 (cyan), 640 (magenta), 1280 (grey).

	ℓ-MVR					e-MVR				
α	0	0.5	1	1.5	2	0	0.5	1	1.5	2
$n = 20$	82.4	84.6	84.1	81.3	77.5	82.3	85.3	84.8	81.6	76.7
$n = 40$	86.6	91.5	90.3	84.4	76.8	89.2	93.0	91.1	85.0	76.1
$n = 80$	90.2	96.0	91.5	81.0	70.1	93.0	96.5	91.4	80.8	67.6
$n = 160$	93.3	97.5	88.0	73.4	61.6	95.4	97.4	87.9	72.7	56.7
$n = 320$	95.4	96.8	82.4	65.1	53.8	96.9	96.6	82.2	63.5	46.4
$n = 640$	97.1	94.4	76.1	57.1	46.6	98.0	94.4	75.7	54.9	37.8
$n = 1280$	98.3	91.8	70.1	50.2	40.4	98.8	91.8	69.6	47.5	30.9

Table 5. Ratio (×100) of MVR average confidence interval lengths for β_4 over corresponding OLS counterpart. Confidence intervals constructed with asymptotic standard errors, assuming correct specification of the CMF.

	ℓ-MVR					e-MVR				
α	0	0.5	1	1.5	2	0	0.5	1	1.5	2
$n = 20$	82.4	84.6	84.1	81.3	77.5	82.3	85.3	84.8	81.6	76.7
$n = 40$	86.6	91.5	90.3	84.4	76.8	89.2	93.0	91.1	85.0	76.1
$n = 80$	90.2	96.0	91.5	81.0	70.1	93.0	96.5	91.4	80.8	67.6
$n = 160$	93.3	97.5	88.0	73.4	61.6	95.4	97.4	87.9	72.7	56.7
$n = 320$	95.4	96.8	82.4	65.1	53.8	96.9	96.6	82.2	63.5	46.4
$n = 640$	97.1	94.4	76.1	57.1	46.6	98.0	94.4	75.7	54.9	37.8
$n = 1280$	98.3	91.8	70.1	50.2	40.4	98.8	91.8	69.6	47.5	30.9

Table 6. Ratio (×100) of MVR average confidence interval lengths for β_4 over corresponding OLS counterpart. Confidence intervals constructed with standard errors with HC1 correction, assuming correct specification of the CMF.

	ℓ-MVR					e-MVR				
α	0	0.5	1	1.5	2	0	0.5	1	1.5	2
$n = 20$	88.9	85.7	79.7	72.5	65.9	87.3	85.0	79.5	72.3	64.6
$n = 40$	93.1	88.8	80.1	69.5	59.9	93.3	88.9	80.1	69.3	58.0
$n = 80$	95.3	90.3	79.0	65.7	54.8	95.7	89.9	78.3	64.4	50.8
$n = 160$	96.7	90.5	76.1	60.4	49.3	97.1	89.9	75.1	58.5	43.6
$n = 320$	97.7	89.9	72.4	54.8	44.2	98.0	89.3	71.3	52.4	36.8
$n = 640$	98.3	88.9	68.5	49.7	39.8	98.6	88.4	67.5	47.1	31.5
$n = 1280$	98.9	87.4	64.4	44.8	35.5	99.0	87.1	63.5	41.9	26.6

Table 7. Ratio (×100) of MVR average confidence interval lengths for β_4 over corresponding OLS counterpart. Confidence intervals constructed with standard errors with HC3 correction, assuming correct specification of the CMF.

	ℓ-MVR					e-MVR				
α	0	0.5	1	1.5	2	0	0.5	1	1.5	2
$n = 20$	99.0	98.7	98.5	97.9	96.4	97.7	100.3	103.1	104.9	104.5
$n = 40$	100.5	100.4	101.4	101.4	98.7	96.8	100.7	105.5	109.4	108.3
$n = 80$	101.8	101.1	102.9	101.8	96.5	96.7	100.8	105.4	108.2	103.2
$n = 160$	102.2	101.3	103.8	99.9	93.1	97.2	100.8	102.7	102.0	92.2
$n = 320$	102.2	101.5	104.5	97.5	90.2	97.7	100.3	98.8	94.1	80.2
$n = 640$	101.9	101.9	104.9	94.8	86.8	98.4	99.1	94.1	85.9	69.5
$n = 1280$	101.3	102.3	104.9	92.5	83.5	98.9	97.7	89.6	78.8	60.6

Table 8. Ratio (×100) of MVR average confidence interval lengths for β_4 over corresponding WLS counterpart. Confidence intervals constructed with asymptotic standard errors, assuming correct specification of the CMF.

	ℓ-MVR					e-MVR				
α	0	0.5	1	1.5	2	0	0.5	1	1.5	2
$n = 20$	99.0	98.7	98.5	97.9	96.4	97.7	100.3	103.1	104.9	104.5
$n = 40$	100.5	100.4	101.4	101.4	98.7	96.8	100.7	105.5	109.4	108.3
$n = 80$	101.8	101.1	102.9	101.8	96.5	96.7	100.8	105.4	108.2	103.2
$n = 160$	102.2	101.3	103.8	99.9	93.1	97.2	100.8	102.7	102.0	92.2
$n = 320$	102.2	101.5	104.5	97.5	90.2	97.7	100.3	98.8	94.1	80.2
$n = 640$	101.9	101.9	104.9	94.8	86.8	98.4	99.1	94.1	85.9	69.5
$n = 1280$	101.3	102.3	104.9	92.5	83.5	98.9	97.7	89.6	78.8	60.6

Table 9. Ratio (×100) of MVR average confidence interval lengths for β_4 over corresponding WLS counterpart. Confidence intervals constructed with standard errors with HC1 correction, assuming correct specification of the CMF.

	ℓ-MVR					e-MVR				
α	0	0.5	1	1.5	2	0	0.5	1	1.5	2
$n = 20$	95.8	97.1	96.9	94.5	90.9	98.8	101.3	103.2	102.7	98.5
$n = 40$	96.3	99.6	100.1	97.0	91.2	99.0	101.0	102.4	101.1	94.3
$n = 80$	97.2	101.5	102.7	97.9	90.1	98.6	99.2	99.5	96.9	87.3
$n = 160$	97.8	102.7	104.5	97.0	88.1	98.5	97.5	95.5	90.6	78.2
$n = 320$	98.4	103.3	105.2	95.5	86.5	98.6	96.7	92.2	84.7	69.4
$n = 640$	98.9	103.3	105.4	93.3	83.4	98.9	95.8	88.8	78.7	61.5
$n = 1280$	99.1	103.2	105.3	91.2	80.0	99.1	94.9	85.6	73.6	55.2

Table 10. Ratio (×100) of MVR average confidence interval lengths for β_4 over corresponding WLS counterpart. Confidence intervals constructed with standard errors with HC3 correction, assuming correct specification of the CMF.

3. Reversal of Fortune: Additional Results

We complement the analysis for the reversal of fortune empirical application in the main text by reporting standard errors with finite-sample corrections (HC1 and HC3 as described in Section 2.1) and results for WLS in Table 11. The exponential scale specification, e-WLS, cannot be used in this empirical application due to several regressors taking value zero for some observations, so that the log transformation cannot be applied to those regressors. Thus we only report the results for WLS with linear scale, ℓ-WLS.

The results in Table 11 further strengthen the main conclusions in the main text with all standard errors increasing slightly when finite-sample corrections are applied, except for MVR when the Americas are dropped (Panel (3)) where the urbanization in 1500 coefficient remains insignificant. For WLS we find that the magnitude of WLS coefficients is smaller than MVR point estimates (except for Panel (3)). In addition to specifications (3), (4), (6) and (9), specification (5) is also found to be not statistically significant with WLS, due to a large drop in the coefficient estimated value relative to both OLS and MVR.

We also report MVR standard errors assuming correct specification of the CMF with and without finite-sample corrections in Table 12. The results confirm that in this example MVR standard errors are robust to finite-sample corrections, and standard errors assuming correct specification of the CMF tend to be slightly smaller.

Overall, we find that our main qualitative conclusions are robust to implementing finite-sample corrections and assuming that the CMF is correctly specified in the calculation of standard errors. Although the numerical simulations in Section 5 of the main text and Section 2 suggest some caution in using ℓ-WLS inference in such small samples, the results in Tables 11-12 provide additional evidence that the relationship between urbanization in 1500 and GDP per capita in 1995 (PPP basis) is weaker and less robust that found using OLS.

	Dependent variable is log GDP per capita (PPP) in 1995											
	OLS	ℓ-WLS	ℓ-MVR	e-MVR	OLS	ℓ-WLS	ℓ-MVR	e-MVR	OLS	ℓ-WLS	ℓ-MVR	e-MVR
	(1) Base sample ($n = 41$)				(2) Without North Africa ($n = 37$)				(3) Without the Americas ($n = 17$)			
Urban. 1500	-0.078	-0.063	-0.067	-0.069	-0.101	-0.097	-0.099	-0.099	-0.115	-0.090	-0.064	-0.077
HC0	0.023	0.021	0.028	0.026	0.032	0.033	0.034	0.034	0.043	0.040	0.127	0.113
HC1	0.023	0.022	0.028	0.027	0.033	0.034	0.035	0.035	0.046	0.043	0.097	0.100
HC3	0.025	0.023	0.029	0.028	0.036	0.037	0.038	0.038	0.056	0.053	0.063	0.082
	(4) Just the Americas ($n = 24$)				(5) With the continent dummies ($n = 41$)				(6) Without neo-Europes ($n = 37$)			
Urban. 1500	-0.053	-0.029	-0.045	-0.044	-0.082	-0.034	-0.063	-0.060	-0.046	-0.034	-0.036	-0.038
HC0	0.029	0.032	0.032	0.032	0.031	0.033	0.029	0.030	0.021	0.020	0.023	0.023
HC1	0.030	0.034	0.033	0.033	0.033	0.035	0.028	0.031	0.022	0.021	0.024	0.024
HC3	0.033	0.044	0.038	0.037	0.035	0.051	0.030	0.029	0.023	0.022	0.025	0.025
	(7) Controlling for Latitude ($n = 41$)				(8) Controlling for colonial origin ($n = 41$)				(9) Controlling for religion ($n = 41$)			
Urban. 1500	-0.072	-0.067	-0.069	-0.070	-0.071	-0.056	-0.063	-0.062	-0.060	-0.015	-0.042	-0.040
HC0	0.020	0.017	0.022	0.021	0.025	0.022	0.026	0.027	0.027	0.029	0.029	0.029
HC1	0.021	0.018	0.022	0.021	0.026	0.023	0.027	0.028	0.029	0.031	0.031	0.030
HC3	0.022	0.019	0.023	0.022	0.028	0.026	0.028	0.029	0.032	0.042	0.034	0.034

TABLE 11. Reversal of fortune. Asymptotic heteroskedasticity-robust OLS, ℓ-WLS and MVR standard errors (HC0), and with finite-sample corrections (HC1 and HC3).

	Dependent variable is log GDP per capita (PPP) in 1995											
	OLS	ℓ-WLS	ℓ-MVR	e-MVR	OLS	ℓ-WLS	ℓ-MVR	e-MVR	OLS	ℓ-WLS	ℓ-MVR	e-MVR
	(1) Base sample ($n = 41$)				(2) Without North Africa ($n = 37$)				(3) Without the Americas ($n = 17$)			
Urban. 1500	-0.078	-0.063	-0.067	-0.069	-0.101	-0.097	-0.099	-0.099	-0.115	-0.090	-0.064	-0.077
HC0	0.023	0.021	0.022	0.022	0.032	0.033	0.033	0.033	0.043	0.040	0.035	0.039
HC1	0.023	0.022	0.022	0.023	0.033	0.034	0.034	0.034	0.046	0.043	0.038	0.041
HC3	0.025	0.023	0.024	0.024	0.036	0.037	0.037	0.037	0.056	0.053	0.051	0.051
	(4) Just the Americas ($n = 24$)				(5) With the continent dummies ($n = 41$)				(6) Without neo-Europes ($n = 37$)			
Urban. 1500	-0.053	-0.029	-0.045	-0.044	-0.082	-0.034	-0.063	-0.060	-0.046	-0.034	-0.036	-0.038
HC0	0.029	0.032	0.030	0.030	0.031	0.033	0.025	0.023	0.021	0.020	0.020	0.021
HC1	0.030	0.034	0.031	0.031	0.033	0.035	0.027	0.025	0.022	0.021	0.021	0.021
HC3	0.033	0.044	0.036	0.035	0.035	0.051	0.029	0.027	0.023	0.022	0.022	0.023
	(7) Controlling for Latitude ($n = 41$)				(8) Controlling for colonial origin ($n = 41$)				(9) Controlling for religion ($n = 41$)			
Urban. 1500	-0.072	-0.067	-0.069	-0.070	-0.071	-0.056	-0.063	-0.062	-0.060	-0.015	-0.042	-0.040
HC0	0.020	0.017	0.018	0.019	0.025	0.022	0.021	0.022	0.027	0.029	0.025	0.026
HC1	0.021	0.018	0.019	0.020	0.026	0.023	0.023	0.023	0.029	0.031	0.027	0.028
HC3	0.022	0.019	0.020	0.021	0.028	0.026	0.025	0.025	0.032	0.042	0.030	0.031

TABLE 12. Reversal of fortune. Asymptotic heteroskedasticity-robust OLS, ℓ-WLS and MVR standard errors assuming correct specification of the CMF (HC0), and with finite-sample corrections (HC1 and HC3).

4. Demand for Gasoline in the United States

4.1. **Empirical Application.** To illustrate our methods further, we consider a second empirical application to the parametric approximation of demand for gasoline in the United States. We use the same data set as in Blundell, Horowitz and Parey (2012), which comes from the 2001 National Household Travel Survey, conducted between March 2001 and May 2002[1]. Blundell, Horowitz and Parey (2012) perform both parametric and nonparametric estimation of the average demand function, and provide evidence of nonlinearities. The data set for their main specifications is large, with a sample of 5254 individual households, and contains household level variables, including gasoline price and consumption, and demographic characteristics. We use these features of the data set to compare the approximation properties of MVR and OLS , to implement our inference methods under misspecification and to calibrate our numerical simulations.

We consider an MVR approximation for the demand for gasoline function

$$Y = \beta_0 + X_1\beta_1 + X_2\beta_2 + X_3'\beta_3 + s(\gamma_0 + X_1\gamma_1 + X_2\gamma_2 + X_3'\gamma_3)e,$$

where e satisfies the orthogonality conditions $E[Xe] = 0$ and $E[Xs_1(X'\gamma)(e^2-1)] = 0$, with $X = (1, X_1, X_2, X_3')'$ and $\gamma = (\gamma_0, \gamma_1, \gamma_2, \gamma_3')'$. We take the outcome Y to be log gasoline annual consumption in gallons, X_1 is log average price in dollars per gallon in county of residence, and X_2 is log income in dollars with each household assigned to 1 of 18 income groups. Following Blundell, Horowitz and Parey (2012), the baseline specification only includes log price and log income, and further covariates are added in other specifications. The vector of additional controls X_3 includes the log of age of household respondent, household size, number of drivers and workers in the household (specification (2)), as well as a dummy for public transport availability (specification (3)), 4 urbanity dummies (specification (4)), 8 population density dummies and 9 regional dummies (specification (5)).

Table 13 reports estimates and standard errors for the average price and income elasticities obtained by OLS, ℓ-MVR and e-MVR across the 5 linear specifications. In the baseline specification, MVR price elasticities are -0.89 and exactly coincide with the average price elasticity found by Yatchew and No (2001) and West (2004), and differ slightly from the OLS point estimate -0.93 in this sample. For specifications (1)-(4), MVR price elasticities are slightly smaller than OLS estimates, and the price

[1]See Blundell, Horowitz and Parey (2012) and ONRL (2004) for a detailed description of the data.

Dependent variable is log of annual household gasoline demand in gallons					
Log price coefficient $\hat{\beta}_1$			Log income coefficient $\hat{\beta}_2$		
OLS	ℓ-MVR	e-MVR	OLS	ℓ-MVR	e-MVR
(1) Baseline specification					
-0.925	-0.892	-0.888	0.289	0.283	0.283
(0.150)	(0.144)	(0.144)	(0.0190)	(0.0173)	(0.0172)
(2) With demographics					
-0.879	-0.857	-0.854	0.246	0.244	0.244
(0.143)	(0.137)	(0.137)	(0.0183)	(0.0169)	(0.0167)
(3) With demographics and public transports					
-0.830	-0.820	-0.816	0.269	0.268	0.268
(0.143)	(0.137)	(0.137)	(0.0187)	(0.0172)	(0.0171)
(4) With demographics, public transports and urbanity					
-0.495	-0.483	-0.478	0.298	0.301	0.301
(0.141)	(0.135)	(0.134)	(0.0190)	(0.0174)	(0.0173)
(5) With demographics, public transports, urbanity and regions					
-0.358	-0.415	-0.408	0.297	0.302	0.302
(0.270)	(0.256)	(0.256)	(0.0199)	(0.0181)	(0.0181)

TABLE 13. Demand for gasoline. Asymptotic heteroskedasticity-robust OLS standard errors and MVR standard errors are in parenthesis.

elasticity drops sharply in specification (4) which adds indicators for urbanity and population density. Adding regional dummies (Panel (5)) results in a further reduction in price elasticities and a loss of significance, although to a much smaller extent for MVR estimates[2]. Given the large sample size, it is interesting to note that for all specifications MVR and OLS standard errors still differ, with MVR standard errors

[2]The p-values for price elasticities increase to 0.185 for OLS and to 0.105 and 0.111 for MVR estimates.

smaller than heteroskedasticity-corrected OLS standard errors, which is a reflection of the heteroskedasticity detected for all specifications[3].

4.2. **Numerical Simulations.** We assess and illustrate the finite-sample properties of our estimators in a Monte Carlo experiment calibrated to our second empirical example. Our models feature a linear CMF, and we implement OLS and MVR with linear and exponential scale functions.

The explanatory variables included in the simulations are chosen according to specification (4) in the demand for gasoline example, the preferred linear specification in Blundell, Horowitz and Parey (2012) (the log price coefficient is no longer significant in specification (5)). We report estimation and inference simulation results for log price and log income, but include all covariates in the simulations. All designs are calibrated to specification (4) by Gaussian maximum likelihood.

Design LOC. Our first design is the homoskedastic model

$$Y = \beta_0 + X_1\beta_1 + X_2\beta_2 + X_3'\beta_3 + \sigma\varepsilon, \quad \varepsilon \sim \mathcal{N}(0,1).$$

Design LIN. Our second design is a set of heteroskedastic models with linear-polynomial scale functions

$$Y = \beta_0 + X_1\beta_1 + X_2\beta_2 + X_3'\beta_3 + (X'\gamma)^\alpha\varepsilon, \quad \varepsilon \sim \mathcal{N}(0,1), \quad \alpha \in \{0.5, 1, 1.5, 2\}.$$

Design EXP. Our third design is a set of heteroskedastic models with exponential-polynomial scale functions

$$Y = \beta_0 + X_1\beta_1 + X_2\beta_2 + X_3'\beta_3 + \exp(X'\gamma)^\alpha\varepsilon, \quad \varepsilon \sim \mathcal{N}(0,1), \quad \alpha \in \{0.5, 1, 1.5, 2\}.$$

For all experiments, we set the sample size to $n = 500, 1000$, and 5254, the sample size in the empirical application, and 5000 simulations are performed. For $n = 5254$, we fix X to the values in the data set, whereas for the smaller sample sizes we draw X with replacement from the values in the data set and keep them fixed across replications. The location design LOC serves as a benchmark for comparing the relative performance of MVR and OLS when OLS is efficient. For $\alpha = 1$, ℓ-MVR is correctly specified for the design LIN, and e-MVR is correctly specified for design EXP. Designs with $\alpha = 0.5$ feature low heteroskedasticity, whereas $\alpha = 2$ corresponds to high heteroskedasticity.

[3]For each specification we implemented the tests of Breusch and Pagan (1979), White (1980) and Koenker (1981)) for heteroskedasticity for OLS and the MVR-based test introduced in Section 4 of the main text. All tests reject the null of homoskedasticity for all specifications.

	Design	LOC	LIN				EXP			
	α	0	0.5	1	1.5	2	0.5	1	1.5	2
		Log price coefficient $\hat{\beta}_1$								
	$n = 500$	102.1	100.5	96.0	89.0	80.6	100.6	96.1	89.0	80.3
ℓ-MVR	$n = 1000$	100.5	98.7	93.6	85.4	75.0	98.8	93.6	85.3	75.2
	$n = 5254$	100.1	98.5	93.4	85.2	74.5	98.5	93.4	84.9	73.8
	$n = 500$	102.1	100.4	95.9	89.1	80.5	100.4	95.8	88.7	79.5
e-MVR	$n = 1000$	100.5	98.7	93.6	85.9	76.4	98.7	93.5	85.3	74.8
	$n = 5254$	100.1	98.5	93.6	86.0	76.5	98.5	93.4	85.1	74.3
		Log income coefficient $\hat{\beta}_2$								
	$n = 500$	101.6	99.6	93.7	85.1	75.1	99.4	92.8	82.9	71.2
ℓ-MVR	$n = 1000$	101.4	98.2	89.4	77.0	63.5	97.4	86.0	70.3	53.9
	$n = 5254$	100.3	96.8	88.1	76.2	63.2	96.0	85.2	70.6	55.3
	$n = 500$	101.7	99.8	93.8	84.8	73.9	99.6	92.6	81.9	68.8
e-MVR	$n = 1000$	101.4	98.3	89.1	75.8	61.1	97.3	84.9	67.4	48.9
	$n = 5254$	100.3	96.7	87.6	74.9	61.0	95.8	84.0	67.7	50.4

TABLE 14. Ratio ($\times 100$) of MVR RMSE for β_1 and β_2 over corresponding OLS counterpart.

Table 14 reports a first set of results regarding the accuracy of our estimators. We report the ratios of RMSEs for β_1 and β_2 of ℓ-MVR and e-MVR over RMSEs of OLS, in percentage terms. The results show that MVR estimators achieve large gains relative to OLS in the presence of heteroskedasticity, with ratios that reach 73.8 for $\hat{\beta}_1$ and 50.4 for $\hat{\beta}_2$ under heteroskedasticity, with e-MVR outperforming ℓ-MVR slightly in this example. Gains in estimation precision increase with the degree of heteroskedasticity and sample size. In the homoskedastic case where OLS is efficient, there is close to no loss in precision from using MVR, with ratios ranging from 100.1 to 102.1. OLS and MVR become equivalent as sample size increases for the homoskedastic case.

Table 15 reports ratios of ℓ-MVR and e-MVR average confidence interval lengths across simulations for β_1 and β_2 over OLS average confidence interval lengths, in percentage terms. In these simulations MVR yields substantially tighter confidence

	Design	LOC	LIN				EXP			
	α	0	0.5	1	1.5	2	0.5	1	1.5	2
		Log price coefficient $\hat{\beta}_1$								
	$n = 500$	98.8	98.1	96.1	92.8	88.6	98.1	96.1	92.8	88.3
ℓ-MVR	$n = 1000$	99.2	98.3	95.8	91.6	86.0	98.3	95.8	91.6	85.7
	$n = 5254$	99.8	98.9	96.3	91.9	86.0	98.9	96.3	91.8	85.6
	$n = 500$	98.6	97.9	95.8	92.5	88.1	97.9	95.8	92.3	87.5
e-MVR	$n = 1000$	99.1	98.3	95.8	92.0	86.9	98.3	95.8	91.6	85.9
	$n = 5254$	99.8	98.9	96.4	92.4	87.2	98.9	96.3	91.9	85.9
		Log income coefficient $\hat{\beta}_2$								
	$n = 500$	98.9	98.2	95.5	91.3	86.3	98.0	95.0	90.1	84.0
ℓ-MVR	$n = 1000$	99.2	98.0	93.9	87.7	80.1	97.5	92.3	84.1	74.4
	$n = 5254$	99.8	98.3	93.9	87.5	79.9	97.9	92.4	84.3	74.7
	$n = 500$	98.7	97.8	94.9	90.4	84.6	97.7	94.3	88.8	81.6
e-MVR	$n = 1000$	99.1	97.8	93.5	86.8	78.5	97.3	91.5	82.2	70.8
	$n = 5254$	99.8	98.2	93.6	86.8	78.4	97.8	91.7	82.5	71.4

TABLE 15. Ratio ($\times 100$) of MVR average confidence interval lengths for β_1 and β_2 over corresponding OLS counterpart. Confidence intervals constructed with asymptotic standard errors.

intervals compared to OLS in the presence of heteroskedasticity, with confidence interval lengths ratios that reach 78.4 for $\hat{\beta}_1$ and 70.8 for $\hat{\beta}_2$, while not incurring any loss in precision for the homoskedastic data generating process. The relative performance of e-MVR improves with the degree of heteroskedasticity.

For completeness we also report results for confidence intervals constructed assuming correct specification of the CMF. Table 16 reports ratios of ℓ-MVR and e-MVR average confidence interval lengths across simulations for β_1 and β_2 over OLS average confidence interval lengths, in percentage terms. MVR confidence intervals are slightly more favorable to MVR compared to the results obtained with standard errors robust to mean misspecification reported in Table 15, while not incurring any loss in precision for the homoskedastic data generating process.

	Design	LOC	LIN				EXP			
	α	0	0.5	1	1.5	2	0.5	1	1.5	2
		Log price coefficient $\hat{\beta}_1$								
	$n = 500$	96.8	96.1	93.9	90.5	86.0	96.1	94.0	90.5	85.8
ℓ-MVR	$n = 1000$	98.4	97.6	95.0	90.7	84.7	97.6	95.0	90.6	84.5
	$n = 5254$	99.7	98.8	96.2	91.8	85.8	98.8	96.2	91.7	85.4
	$n = 500$	97.4	96.3	94.2	91.0	86.7	96.3	94.2	90.8	86.1
e-MVR	$n = 1000$	98.6	97.6	95.2	91.4	86.4	97.6	95.2	91.1	85.4
	$n = 5254$	99.7	98.8	96.3	92.4	87.1	98.8	96.2	91.9	85.8
		Log income coefficient $\hat{\beta}_2$								
	$n = 500$	96.5	95.8	93.1	88.9	83.7	95.7	92.7	87.8	81.8
ℓ-MVR	$n = 1000$	98.1	97.1	93.2	87.0	79.5	96.7	91.6	83.6	74.0
	$n = 5254$	99.7	98.2	93.8	87.4	79.8	97.8	92.3	84.2	74.7
	$n = 500$	97.1	95.8	92.9	88.4	82.7	95.6	92.3	86.9	79.9
e-MVR	$n = 1000$	98.6	97.0	92.7	86.0	77.8	96.5	90.7	81.6	70.4
	$n = 5254$	99.7	98.1	93.5	86.7	78.4	97.7	91.7	82.5	71.4

TABLE 16. Ratio ($\times 100$) of MVR average confidence interval lengths for β_1 and β_2 over corresponding OLS counterpart. Confidence intervals constructed with asymptotic standard errors assuming correct specification of the CMF.

4.3. **Additional Simulations: Nonlinear CMF.** We present the results of a second set of experiments in which we compare the approximation properties of MVR to those of OLS under misspecification of the CMF, in RMSE. The designs of our simulations are modified to incorporate a nonlinear relationship between X_1 (log price) and Y (log gasoline annual consumption). We specify the nonlinear relationship in X_1 by means of trigonometric basis functions

$$f(x_1, \delta_1) = \delta_{11} x_1 + \delta_{12} \sin(2\pi x_1) + \delta_{13} \cos(2\pi x_1) + \delta_{14} \sin(4\pi x_1) + \delta_{15} \cos(4\pi x_1).$$

All designs are calibrated to specification (4) by Gaussian maximum likelihood.

Design LOC. Our first design is the homoskedastic model

$$Y = \beta_0 + f(X_1, \beta_1) + X_2\beta_2 + X_3'\beta_3 + \sigma\varepsilon.$$

	Design	LOC	LIN				EXP			
	α	0	0.5	1	1.5	2	0.5	1	1.5	2
	$n = 500$	101.2	100.9	100.0	98.5	96.5	100.9	100.0	98.6	96.5
ℓ-MVR	$n = 1000$	100.7	100.0	97.9	94.6	90.1	100.0	97.9	94.1	88.9
	$n = 5254$	100.1	99.6	97.8	95.0	91.6	99.5	97.5	94.1	89.6
	$n = 500$	100.8	100.6	99.9	98.6	96.8	100.6	99.8	98.3	96.2
e-MVR	$n = 1000$	100.6	99.9	97.8	94.5	90.2	99.9	97.7	93.8	88.3
	$n = 5254$	100.1	99.6	97.8	95.0	91.6	99.5	97.4	93.9	89.1

TABLE 17. Ratio ($\times 100$) of average MVR RMSE for $\mu(x)$ over corresponding OLS counterpart.

Design LIN. Our second design is a set of heteroskedastic models with linear-polynomial scale functions

$$Y = \beta_0 + f(X_1, \beta_1) + X_2\beta_2 + X_3'\beta_3 + s(\gamma_0 + f(X_1, \gamma_1) + X_2\gamma_2 + X_3'\gamma_3)^\alpha \varepsilon, \quad \alpha \in \{0.5, 1, 1.5, 2\}.$$

Design EXP. Our third design is a set of heteroskedastic models with exponential-polynomial scale functions

$$Y = \beta_0 + f(X_1, \beta_1) + X_2\beta_2 + X_3'\beta_3 + s(\gamma_0 + f(X_1, \gamma_1) + X_2\gamma_2 + X_3'\gamma_3)^\alpha \varepsilon, \quad \alpha \in \{0.5, 1, 1.5, 2\}.$$

where $\varepsilon \sim \mathcal{N}(0, 1)$. For all designs we implement MVR and OLS for the same sample sizes and X values as in Section 4.2, with the number of simulations set to 5000.

Table 17 reports results regarding the accuracy of OLS and MVR linear approximations of the $\mu(x, \beta) = \beta_0 + f(x_1, \beta_1) + x_2\beta_2 + x_3'\beta_3$, evaluated at the n sample values x_{1i} of X_1, and at fixed values of the remaining variables.[4] For each data generating process we report the ratios of average estimation errors across simulations of ℓ-MVR and e-MVR relative to OLS in percentage terms. Estimation errors are measured for each simulation in RMSE, and then averaged across simulations.

In these simulations MVR yields more accurate approximation of nonlinear CMFs than OLS, measured in RMSE. Thus, in presence of heteroskedasticity the minimum

[4]The non binary variables X_2, $X_{31}, \dots X_{34}$, are evaluated at their modal values. These variables are the log of household income, age of household respondent, household size, number of drivers and workers in the household, respectively. We fix the value of the remaining indicators for public transport availability, urbanity and population density included in X_3 to one.

mean squared error OLS property does not necessarily translate into more accurate CMF approximation in finite samples relative to MVR.

References

Blundell, R., Horowitz, J. & Parey, M. (2012). Measuring the Price Responsiveness of Gasoline Demand: Economic Shape Restrictions and Nonparametric Demand Estimation. *Quantitative Economics* 3, pp. 29–51.

Breusch, T. S. and Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica* (47, September), pp. 1287–1294.

Koenker, R. (1981). A Note on Studentizing a Test for Heteroscedasticity. *Journal of Econometrics* 17, pp. 107–112.

Long, J. S. and Ervin, L. H. (2000). Using Heteroscedasticity Consistent Standard Errors in the Linear Regression Model. *The American Statistician* 54, pp. 217–224.

MacKinnon, J. G. (2013). Thirty Years of Heteroskedasticity-Robust Inference. In *Recent advances and future directions in causality, prediction, and specification analysis*, pp. 437-461. Springer, New York, NY.

MacKinnon, J. G. and White, H. (1985). Some Heteroskedasticity-Consistent Covariance Matrix Estimators with Improved Finite Sample Properties. *Journal of Econometrics* 29, pp. 305–325.

ONRL (2004). 2001 National Household Travel Survey. User's Guide, Oak Ridge National Laboratory.

White, H. (1980). A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity. *Econometrica* (48, May), pp. 817–838.

West, S. (2004). Distributional Effects of Alternative Vehicle Pollution Control Policies. *Journal of Public Economics* 88, pp. 735–757.

Yatchew, A. and No, J. A. (2001). Household Gasoline Demand in Canada. *Econometrica* (69, November), pp. 1697–1709.