Moreira, Marcelo J.; Mourão, Rafael; Moreira, Humberto

Working Paper

A critical value function approach, with an application to persistent time-series

cemmap working paper, No. CWP24/16

Provided in Cooperation with:
Institute for Fiscal Studies (IFS), London

Suggested Citation: Moreira, Marcelo J.; Mourão, Rafael; Moreira, Humberto (2016) : A critical value function approach, with an application to persistent time-series, cemmap working paper, No. CWP24/16, Centre for Microdata Methods and Practice (cemmap), London, https://doi.org/10.1920/wp.cem.2016.2416

This Version is available at:
https://hdl.handle.net/10419/149771

cemmap
centre for microdata methods and practice

Supplement for "A Critical Value Function Approach, with an Application to Persistent Time-Series"

Marcelo J. Moreira
Rafael Mourão
Humberto Moreira

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP24/16

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL
An ESRC Research Centre

Supplement to "A Critical Value Function Approach, with an Application to Persistent Time-Series"

Marcelo J. Moreira, Rafael Mourão, and Humberto Moreira

Escola de Pós-Graduação em Economia e Finanças (FGV/EPGE)
Getulio Vargas Foundation - 11th floor
Praia de Botafogo 190
Rio de Janeiro - RJ 22250-040
e-mail: moreiramj@fgv.br

1. Introduction

This supplement consists of two parts. The first section includes additional graphs on the critical value function, size and power comparison. The second section provides additional proofs for the remaining theoretical results.

2. Numerical Results

We provide additional numerical results. First, we allow for a time trend and compare the performance of commonly-used sampling methods to our CVF approach. We then consider flattening schemes to guarantee that the critical value function (CVF) converges to the usual standard normal quantiles. Finally, we make further power comparisons between our similar t-test and other similar tests for different values of the autoregressive parameter γ and degrees of endogeneity ρ.

2.1. Time Trend

For now, we extend the model to allow for a time trend. The model is given by

$$\begin{aligned} y_t &= \mu_y' d_t + \beta x_{t-1} + \epsilon_t^y \\ x_t &= \gamma x_{t-1} + \epsilon_t^x, \end{aligned}$$

where μ_y is a two-dimensional vector and $d_t = (1, t)'$. We consider the group of translation transformations on the data

$$g \circ (y, x) = (y + Dg, x),$$

where g is a two-dimensional vector and D is a $T \times 2$ matrix whose t-th line is given by d_t'. This yields a transformation on the parameter space

$$g \circ \left(\beta, \gamma, \mu_y\right) = \left(\beta, \gamma, \mu_y + g\right).$$

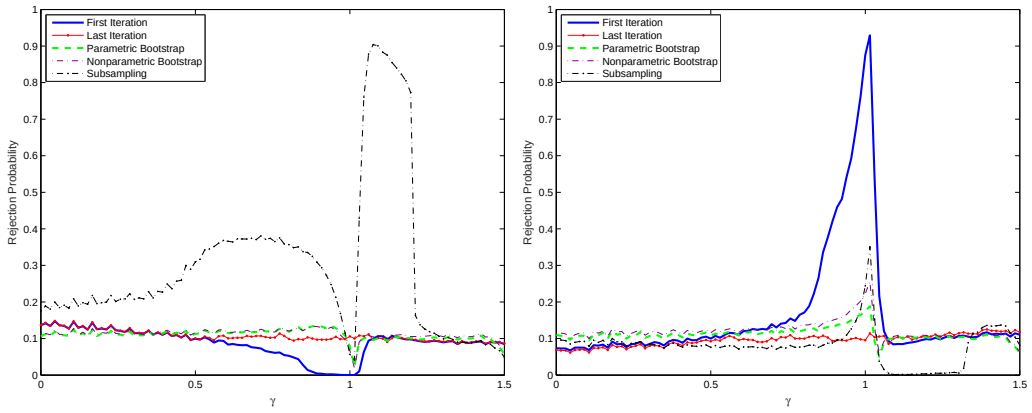
Any invariant test can be written as a function of the maximal invariant statistic. Let $q = (q_1, q_2)$ be an orthogonal $T \times T$ matrix where the first column is given by $q_1 = D\left(D'D\right)^{-1/2}$. Algebraic manipulations show that $q_2 q_2' = M_D$, where $M_D = I_T - D\left(D'D\right)^{-1} D'$ is the projection matrix to the space orthogonal to 1_T. Let x_{-1} be the

T-dimensional vector whose t-th entry is x_{t-1}, and define $w^\mu = q_2' w$ for a T-dimensional vector w. The maximal invariant statistic is given by $r = (y^\mu, x)$. Its density function is given by

$$f_{\beta,\gamma}\left(y^\mu, x\right) = \left(2\pi\sigma_{xx}\right)^{-\frac{T}{2}} \exp\left\{-\frac{1}{2\sigma_{xx}}\sum_{t=1}^{T}\left(x_t - x_{t-1}\gamma\right)^2\right\} \tag{2.1}$$
$$\times \left(2\pi\sigma_{yy.x}\right)^{-\frac{T-2}{2}} \exp\left\{-\frac{1}{2\sigma_{yy.x}}\sum_{t=1}^{T}\left(y_t^\mu - x_t^\mu\frac{\sigma_{xy}}{\sigma_{xx}} - x_{t-1}^\mu\left[\beta - \gamma\frac{\sigma_{xy}}{\sigma_{xx}}\right]\right)^2\right\},$$

where $\sigma_{yy.x} = \sigma_{yy} - \sigma_{xy}^2/\sigma_{xx}$ is the variance of ϵ_t^y not explained by ϵ_t^x. Because of invariance, the density depends on β and γ, but not on μ_y. Furthermore, the formulae for the sufficient statistics and the log-likelihood ratio process are exactly the same as in the model with only an intercept, but replacing q_2 in y^μ and x^μ. The asymptotic distribution of the new $R_T(\gamma)$ and $K_T(\gamma)$ statistics is, however, different when we allow for time trend. In particular, the model curvature increases in the sense of Efron [1, 2] when $\gamma = 1$. This yields larger departures from the nominal size by conventional sampling schemes.

Fig S1: Size (intercept and time trend)

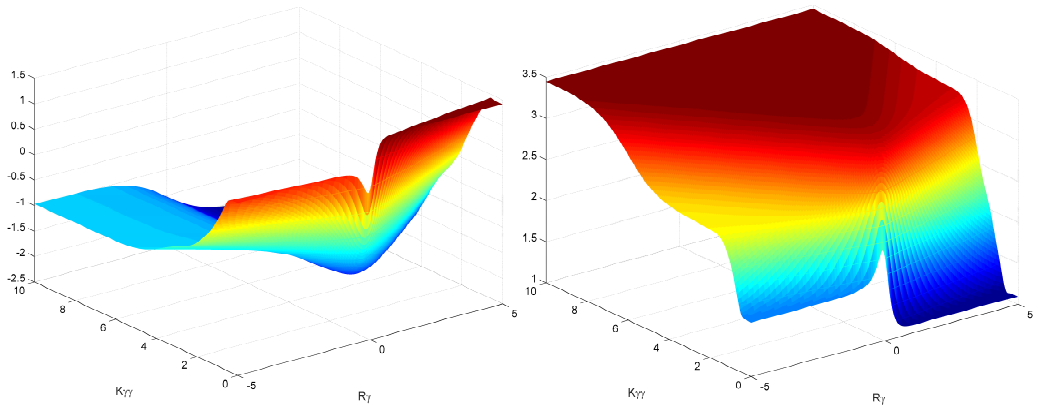


(a) $\rho = 0.95$ (b) $\rho = -0.95$

Figure S1 plots null rejection probabilities when we allow for a time trend. Comparing this graph with Figure 1, it is evident that both bootstrap methods and subsampling have null rejection probabilities farther away from the 10% nominal size when we allow for a time trend. The CVF, however, is able to correct size distortions whether there is a time trend or not.

As such, the CVF needs to provide more curvature when we allow for time trend. This can be seen by comparing Figure S2, which plots the CVF when the model has a time trend, to Figure 2. The range of values that the CVF can take is larger when we make inference in the presence of a time trend. For example, take the CVF plot when $\rho = 0.95$. The CVF can have values as low as -2.5 and as high as 1.5 in the presence of a time trend. On the other hand, the CVF ranges from -1 to 1.5 when we make inference knowing there is no time trend.

Fig S2: Critical Value Function (intercept and time trend)



(a) $\rho = 0.95$ (b) $\rho = -0.95$

2.2. Flattening Schemes

We study the behavior of the critical value function (CVF) when the series is stationary and explosive. We focus on the main model in which there is only intercept and no time trend. We are able to analyze the behavior of $R_\gamma/K_{\gamma\gamma}$ and $K_{\gamma\gamma}$ when the series is stationary and explosive (to save space, we again use the notation R_γ and $K_{\gamma\gamma}$ for $R_{\gamma,T}(1)$ and $K_{\gamma\gamma,T}(1)$, respectively). Figure S3 plots the critical values as a function of $R_\gamma/K_{\gamma\gamma}$ and $K_{\gamma\gamma}$. It is convenient to re-scale the x-axis and y-axis in the figures to analyze the critical value under different asymptotic regimes. We choose the scale

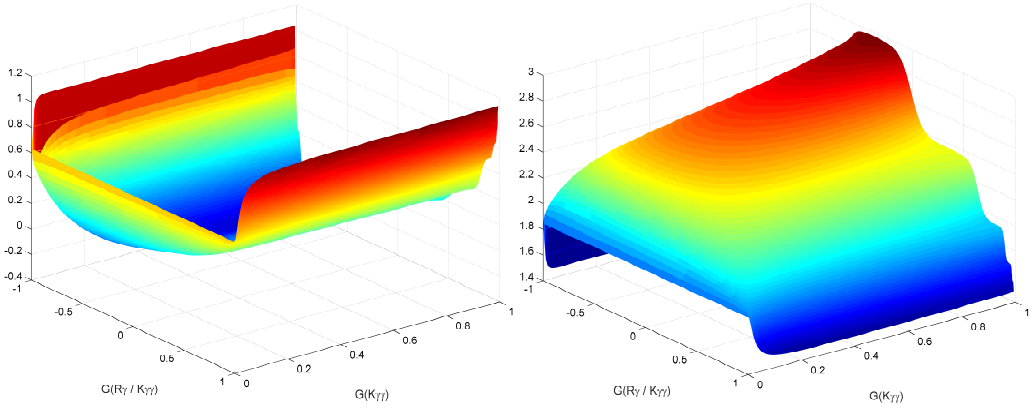
$$G(z) \equiv 2.(F(z) - 0.5),$$

where $F(z) = [1 + \exp(-z)]^{-1}$ is the logistic distribution. This scale is an increasing, symmetric transformation, which preserves zero and maps ∞ into 1; that is, $G'(z) > 0$, $G(-z) = -G(z)$, $G(0) = 0$, and $\lim_{z\to\infty} G(z) = 1$.

When the series is stationary, $R_\gamma/K_{\gamma\gamma}$ converges to $-\infty$ (under local alternatives) and $K_{\gamma\gamma}$ to 0. This convergence is represented by the point $(-1, 0)$ in the domain for $G(R_\gamma/K_{\gamma\gamma})$ and $G(K_{\gamma\gamma})$. When the series is explosive, $R_\gamma/K_{\gamma\gamma}$ diverges to ∞ (under local alternatives) and $K_{\gamma\gamma}$ to ∞. This limit is indicated by the point $(1, 1)$ in the new domain. In both situations, the critical value function takes values near 1.28, the 90% quantile of a standard normal distribution. Of course, those values are not exactly equal to 1.28. This suggests we can try to flatten the critical value function using the values of $R_\gamma/K_{\gamma\gamma}$ or $K_{\gamma\gamma}$ themselves. For example, we could consider setting the critical value function at 1.28 for (i) large values of $|R_\gamma/K_{\gamma\gamma}|$, or (ii) small values and large values of $K_{\gamma\gamma}$. The figures below consider flattening schemes when (i) $|R_\gamma/K_{\gamma\gamma}| > 10^2$ and (ii) $K_{\gamma\gamma} < 10^{-2}$ or $K_{\gamma\gamma} > 10^6$. For these regions, the critical value function is forced to take the value 1.28. Figure S4 plots the null rejection probabilities from both flattening methods for $\rho = 0.95$ and -0.95.

The tests' sizes are sometimes sensitive to the choice of the thresholds for both flattening methods. When flattening slightly improves size, it does not yield a substantial effect on the power of the respective tests. Hence, we report power using the standard critical value function, without any auxiliary flattening method.

Fig S3: Critical Value Function



(a) $\rho = 0.95$ (b) $\rho = -0.95$

2.3. Additional Power Comparison

Figures S5-S8 present power for different levels of γ and ρ. The results further support the similar t-test in terms of size and power. The behavior of the L_2 test is sensitive to the values of c, sometimes behaving as a one-sided and other times as a two-sided test. At the value $b = 0$, the UMPCU test presents null rejection probabilities above the nominal level when $c = 5$, and close to zero when $c = 30$. This problem is possibly due to the inaccuracy of the algorithm used by Jansson and Moreira [3] for integral approximations. Although this problem can be conceivably fixed by using better numerical approximations for $c > 0$, the power of the UMPCU test is considerably below that found by the similar t-test for $c \leq 0$. Overall, the similar t-test has correct size and better power than the competing tests.

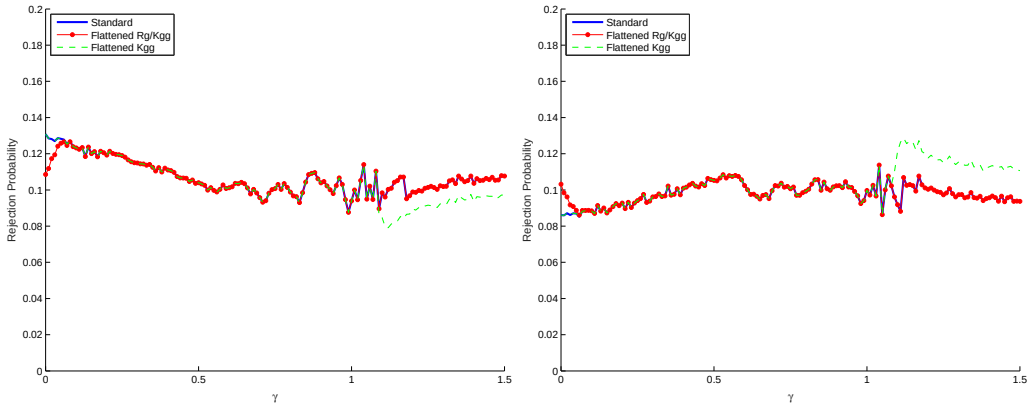
References

[1] Efron, B. (1975). Defining the curvature of a statistical problem (with applications to second order efficiency). *Annals of Statistics 3*, 1189–1242.

[2] Efron, B. (1978). The geometry of exponential families. *Annals of Statistics 6*, 362–376.

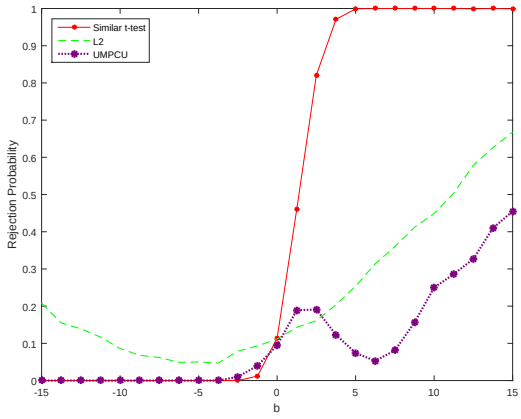
[3] Jansson, M. and M. J. Moreira (2006). Optimal inference in regression models with nearly integrated regressors. *Econometrica 74*, 681–715.

Fig S4: Size

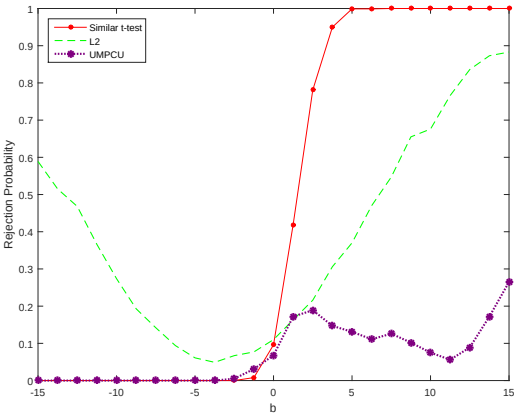


(a) $\rho = 0.95$ (b) $\rho = -0.95$

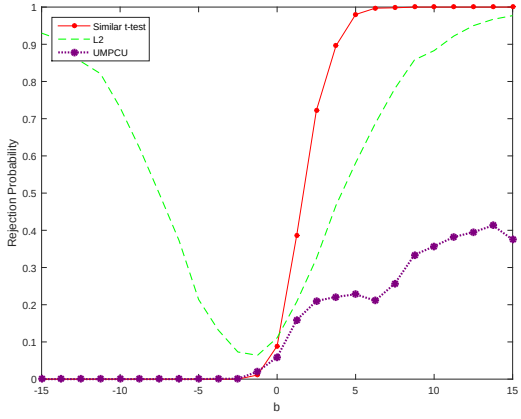
Fig S5: Power ($\rho = 0.5$)



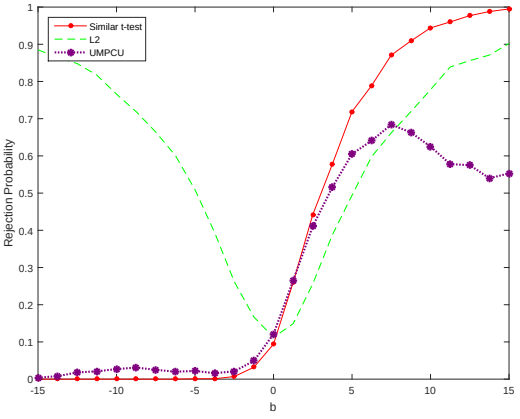
(a) $\gamma = 0$



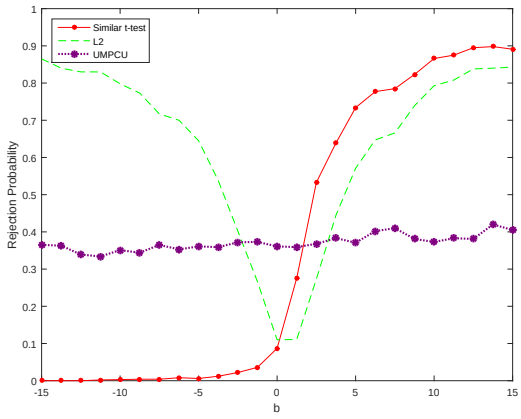
(b) $\gamma = 0.5$



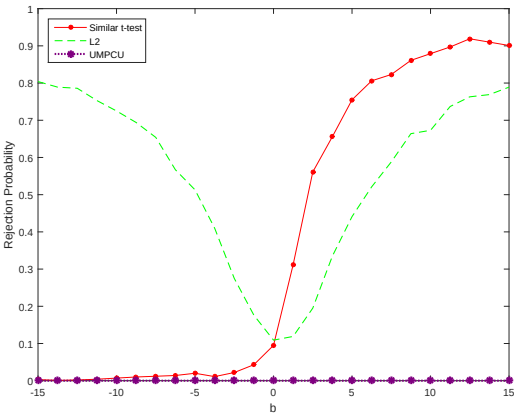
(c) $\gamma = 0.85$



(d) $\gamma = 1.0$

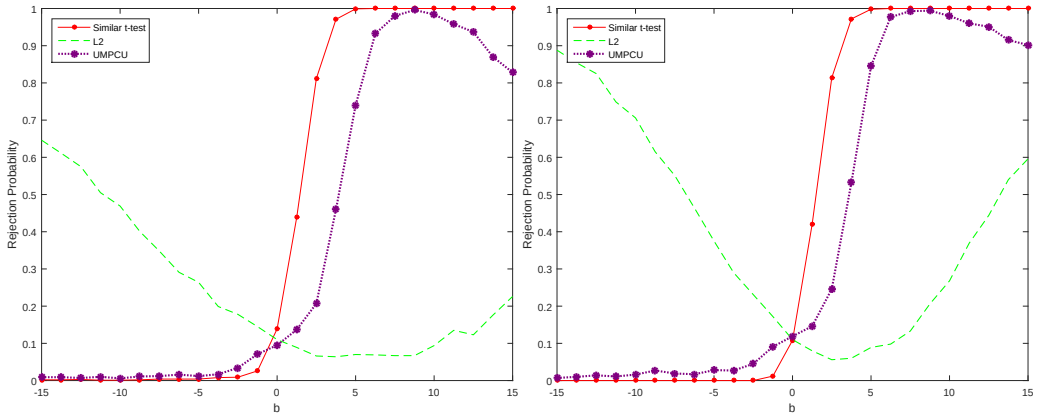


(e) $\gamma = 1.05$

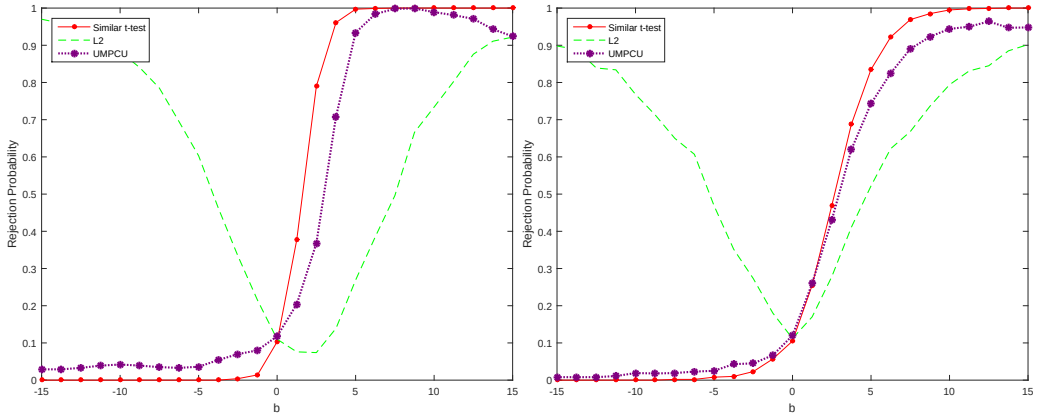


(f) $\gamma = 1.3$

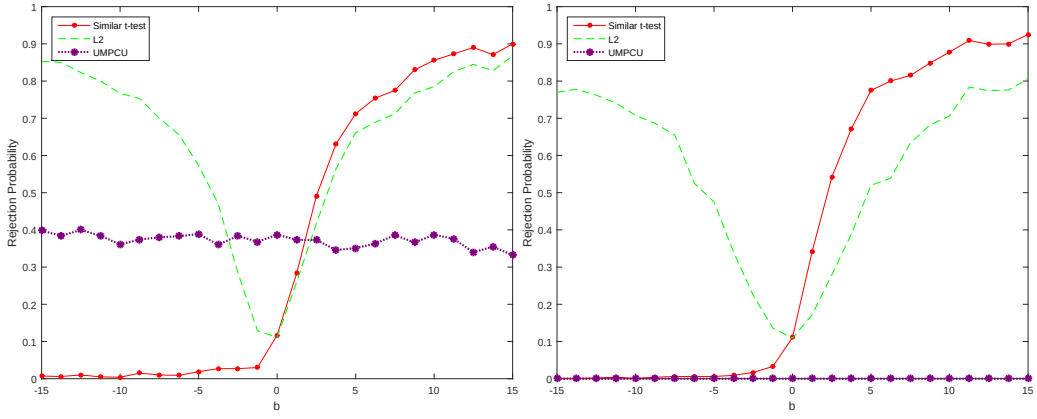
Fig S6: Power ($\rho = -0.5$)



(a) $\gamma = 0$ (b) $\gamma = 0.5$

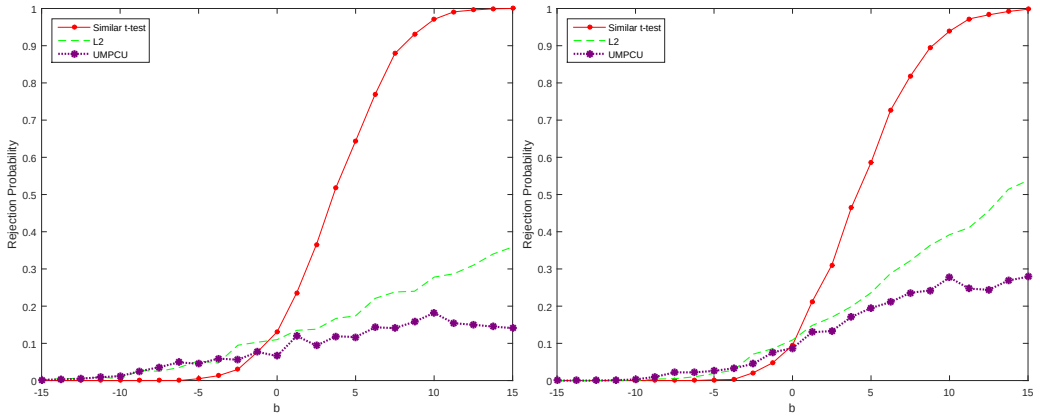


(c) $\gamma = 0.85$ (d) $\gamma = 1.0$

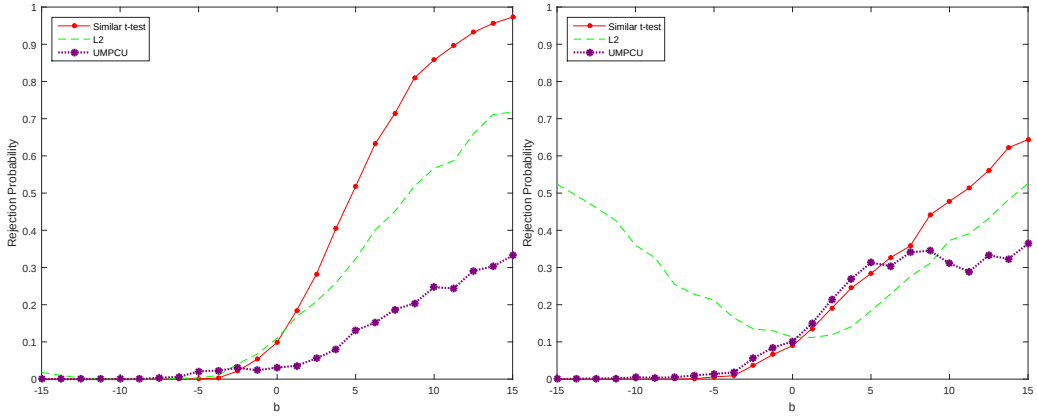


(e) $\gamma = 1.05$ (f) $\gamma = 1.3$

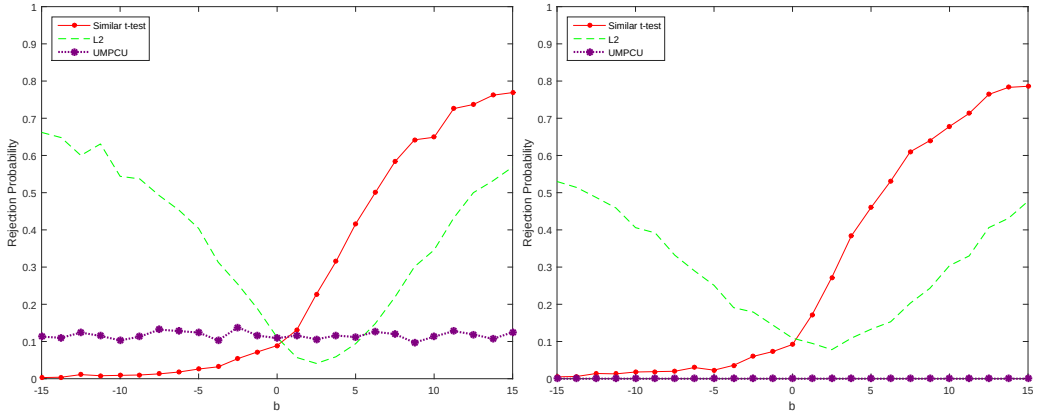
Fig S7: Power ($\rho = 0.95$)



(a) $\gamma = 0$ (b) $\gamma = 0.5$

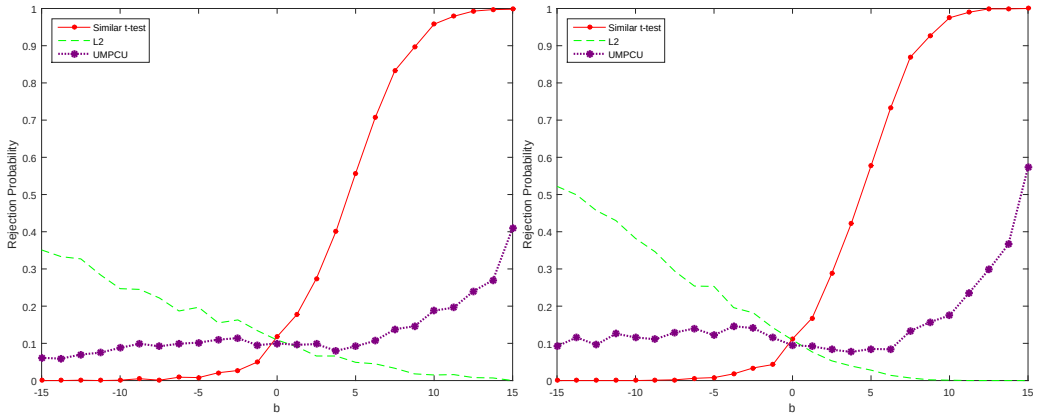


(c) $\gamma = 0.85$ (d) $\gamma = 1.0$

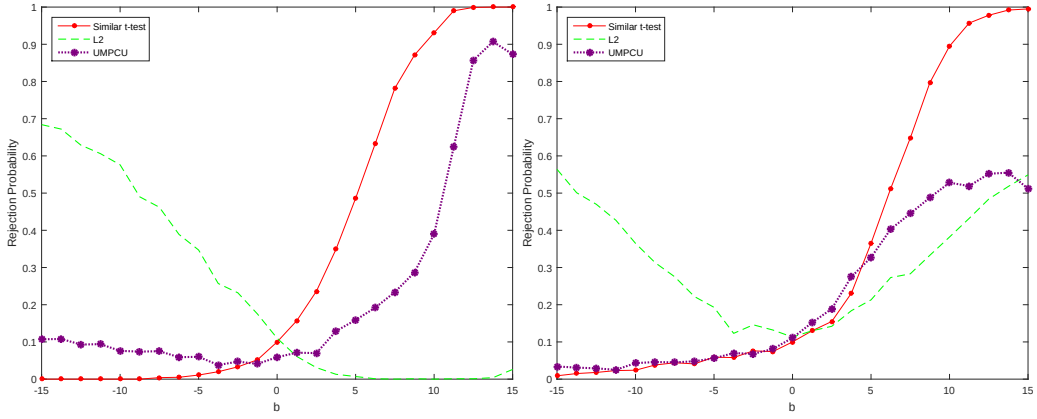


(e) $\gamma = 1.05$ (f) $\gamma = 1.3$

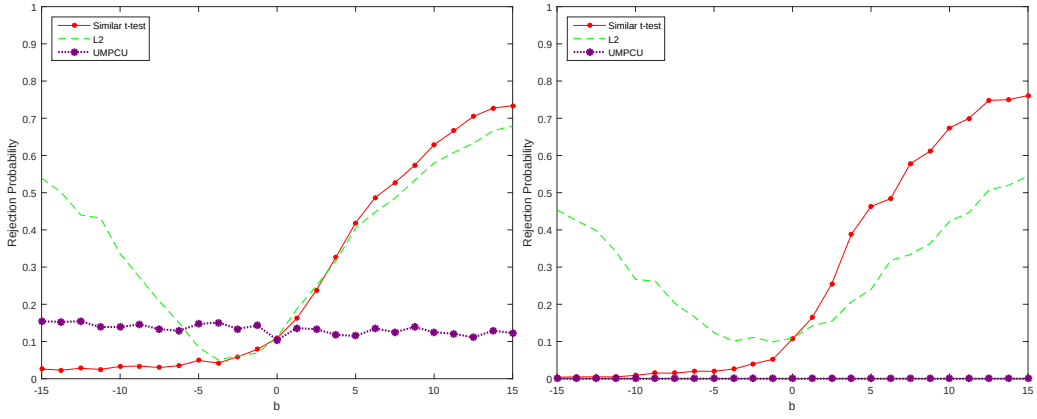
Fig S8: Power ($\rho = -0.95$)



(a) $\gamma = 0$ (b) $\gamma = 0.5$



(c) $\gamma = 0.85$ (d) $\gamma = 1.0$



(e) $\gamma = 1.05$ (f) $\gamma = 1.3$