Further theoretical and practical insight to the do-validated bandwidth selector

ENNO MAMMEN

Department of Economics, University of Mannheim, Germany

emammen@rumms.uni-mannheim.de

MARÍA DOLORES MARTÍNEZ MIRANDA[1]

Cass Business School, City University London, U.K.

University of Granada, Spain

Maria.Miranda.1@city.ac.uk

JENS PERCH NIELSEN

Cass Business School, City University London, U.K.

Jens.Nielsen.1@city.ac.uk

STEFAN SPERLICH

Département des sciences économiques, Universitè de Genève, Switzerland

Stefan.Sperlich@unige.ch

Abstract

Recent contributions to kernel smoothing show that the performance of cross-validated bandwidth selectors improves significantly from indirectness and that the recent do-validated method seems to provide the most practical alternative among these methods. In this paper we show step by step how classical cross-validation improves in theory, as well as in practice, from indirectness and that do-validated estimators improve in theory, but not in practice, from further indirectness. This paper therefore provides a strong support for the practical and theoretical properties of do-validated bandwidth selection. Do-validation is currently being introduced to survival analysis in a number of contexts and this paper provides evidence that this might be the immediate step forward.

[1] *Corresponding author, 106 Bunhill Row, UK - London EC1Y 8TZ, Tel. +44(0)7583643049, e-mails: Maria.Miranda.1@city.ac.uk; mmiranda@ugr.es*

Keywords: Kernel density estimation; Data-adaptive bandwidth selection; Indirect cross-validation; Do-validation.

MSC: 62G07,65D10.

1 Introduction

Indirect cross-validated bandwidth selection has a number of theoretical and practical advantages, see among others Hart and Yi (1998), Hart and Lee (2005), Savchuk et al. (2008, 2010). In this paper classical cross-validation is improved through indirectness considering a series of polynomial kernels as indirect kernels. The limit of this series of polynomial kernels is the popular Gaussian kernel. Asymptotic theory provides evidence that the performance of indirect cross-validation improves with the order of the polynomial kernel with the Gaussian limit being the best performing. This provides some theoretical justification for the fact that the Gaussian kernel is a popular choice in indirect cross-validation. However, we show that even our best performing indirectly cross-validated bandwidth selectors are outperformed by the recent do-validated estimator of Mammen et al. (2011). Therefore we have considered improving do-validation from indirectness using polynomial kernels in the indirect step. Indirect do-validation does indeed outperform classical do-validation from the point of view of asymptotic theory. However, from a practical point of view, do-validation is still the best method. This paper can therefore be seen as an argument in favour of exploring do-validation in kernel smoothing rather than trying to improve it even further. This is indeed being done at the moment, see Gámiz-Pérez et al. (2013a,b,c) for the introduction of do-validation to three fundamental models of survival analysis. Do-validation could be considered in other smoothing problems, see for example Soni et al. (2012), Oliveira et al. (2012), Spreeuw et al. (2013), Lee et al. (2010, 2012a,b), Buch-Kromann and Nielsen (2012), González-Manteiga et al. (2013).

The paper is organized as follows. In Section 2 we first consider indirect cross-validation, where the theoretical and practical improvements of highering the power of the indirect kernel is very clear. Both the theoretical and the finite sample performance improve consistently in every step when increasing the power of the indirect kernel. In Subsection 2.1 we describe a simulation study to assess the finite sample performance of the method. In Section 3 we consider indirect do-validation. The theoretical relative improvements of highering the power of the indirect kernel follow the same pattern as we found in indirect cross-validation. However, the finite sample results are less clear. Section 4 concludes the paper.

2 Indirect cross-validated bandwidth selection in kernel density estimation

In this section we consider indirect cross-validation in its simplest possible version taken from Savchuk et al. (2008, 2010). These papers considered a number of variations of indirect cross-validation. We consider here the simplest possible version, where one has one indirect kernel and one original kernel. The above two papers seem to have some preference for a mixture of Gaussian kernels as indirect kernel. Here we provide a theoretical justification for why this might be a good idea. The Gaussian kernel is in some sense the optimal kernel of a class of indirect kernels, and the theoretical and practical advantage of choosing the Gaussian kernel as indirect kernel can be quite substantial. In our derivation of indirect cross-validation below we use notation similar to the used by Mammen et al. (2011), who considered a class of bandwidth selectors with the indirect cross-validation bandwidth $\widehat{h}_{ICV,L}$, with indirect kernel L, as a special case.

The aim is to get a bandwidth with a small Integrated Squared Error (ISE) for the kernel density estimator

$$\widehat{f}_{h,K}(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{X_i - x}{h}\right),$$

with symmetric kernel function K.

The bandwidth $\widehat{h}_{ICV,L}$ is based on the inspection of the kernel density estimator $\widehat{f}_{h,L}$, for a kernel L that fulfills $L(0) = 0$. And it comes from the following CV score:

$$\int \widehat{f}_{h,L}(x)^2 dx - 2n^{-1}\sum_{i=1}^{n} \widehat{f}_{h,L}(X_i). \qquad (1)$$

Note that because of $L(0) = 0$ we do not need to use a leave-one-out version of $\widehat{f}_{h,L}$ in the sum on the right hand side. Also any kernel L can be defined to fulfill such condition just by setting $L(u) = L(u)1_{u\neq 0}$, and it will be considered hereafter in the indirect step. Thus, the indirect cross-validation bandwidth $\widehat{h}_{ICV,L}$ is defined by

$$\widehat{h}_{ICV,L} = d_L^{-1}\widehat{h}_L$$

with $\widehat{h}_L$ being the minimizer of the score (1). Here, we define

$$d_M = \left(\frac{R(K)}{R(M)}\frac{\mu_2^2(M)}{\mu_2^2(K)}\right)^{-1/5} \tag{2}$$

for a kernel M with $R(g) = \int g^2(x)dx$, $\mu_l(g) = \int x^l g(x)dx$ for functions g and integers $l \geq 0$. Note that the bandwidth $\widehat{h}_L$ is a selector for the density estimator with kernel L. After multiplying with the factor d_L^{-1} it becomes a selector for the density estimator $\widehat{f}_{h,K}$. This follows from classical smoothing theory and has been used at many places in the discussion of bandwidth selectors. Note that the indirect cross-validation method contains the classical cross-validation bandwidth selector as one example with $K = L$.

We now apply results from Mammen et al. (2011) to derive the asymptotic distribution of the difference between the indirect cross-validation bandwidths $\widehat{h}_{ICV,L}$ and the ISE-optimal bandwidth $\widehat{h}_{ISE}$. Here, the bandwidth $\widehat{h}_{ISE}$ is defined by

$$\widehat{h}_{ISE} = \arg\min_h \left[\int \left(\widehat{f}_{h,K}(x) - f(x)\right)^2 dx\right]. \tag{3}$$

Under some mild conditions on the density f and the kernels K and L, see Assumptions (A1) and (A2) in Mammen et al. (2011), one gets by application of their Theorem 1 that

$$n^{3/10}(\widehat{h}_{ICV,L} - \widehat{h}_{ISE}) \to N(0, \sigma^2_{ICV,L}) \qquad \text{in distribution,}$$

with

$$\sigma^2_{ICV,L} = C_{f,K}\left\{4R(K)\frac{\mathrm{V}(f'')}{R(f'')R(f)} + c_{L,K}\right\} \tag{4}$$

where $C_{f,K}$, $R(f'')$, $R(f)$ and $\mathrm{V}(f'')$ are functionals depending on the density f, its derivatives and the kernel K. The functional $c_{L,K}$ is what distinguishes the asymptotic variance of the bandwidth estimates and it only depends on the chosen kernels L and K. See Mammen et al. (2011) for an explicit definition of all of these functionals.

Here we are interested in indirect cross-validation defined with any symmetric kernel K (as for example the Epanechnikov kernel) and as the kernel L a polynomial kernel with higher power. Specifically we define a general kernel function with power r by

$$K_{2r}(u) = \kappa_r (1-u^2)^r 1_{\{-1<u<1\}} \tag{5}$$

with $\kappa_r = (\int_{-1}^{1}(1-u^2)^r du)^{-1}$. Note that for $r = 1$ it is the Epanechnikov kernel and for $r = 2$ it gives the quartic kernel. We now study the theoretical performance of indirect cross-validation for the choice $K = K_2$ (the Epanechnikov kernel) and $L = K_{2r}$ for different choices of r. We start by considering the limiting case $r \to \infty$. For this purpose we consider the kernel $K^*_{2r}(u) = (2r)^{-1/2}K_{2r}((2r)^{-1/2}u)$ that differs from K_{2r} by scale. Because the definition of the bandwidth selector does not depend on the scale of L we have that $\sigma^2_{ICV,K_{2r}} = \sigma^2_{ICV,K^*_{2r}}$. Furthermore, because of $\lim_{r\to\infty}(1-(2r)^{-1}u^2)^r = e^{-u^2/2}$ it holds that, after scaling, the polynomial kernels converge to the Gaussian kernel when r goes to infinity

$$\lim_{r\to\infty}(2r)^{-1/2}K_{2r}((2r)^{-1/2}u) = \phi(u) = \frac{1}{\sqrt{2\pi}}e^{-\frac{u^2}{2}}.$$

Moreover, it holds that $\sigma^2_{ICV,K_{2r}} = \sigma^2_{ICV,K^*_{2r}} \to \sigma^2_{ICV,\phi}$ for $r \to \infty$. This can be shown by dominated convergence using the fact that $(1-(2r)^{-1}u^2)^r \le e^{-u^2/2}$. Thus a Gaussian indirect kernel is a limiting case for the performance of indirect cross-validation.

According to the results above, the asymptotic variance of $\widehat{h}_{ICV,K_{2r}} - \widehat{h}_{ISE}$ is of the form giving in (4) with $c_{L,K} \equiv c_r$ a constant depending on r. We have just argued that $c_r \to c_\infty$ for $r \to \infty$ where $c_\infty = 3.48$ is the constant corresponding to the Gaussian kernel. Figure 1 shows c_r as a function of r. It illustrates the convergence but it also shows that this convergence is monotone: by increasing the power r ($r = 2, 3, 4, \ldots$) we get an incremental reduction in the asymptotic variance factor for indirect cross-validation.

One sees that the trick of indirect cross-validation significantly improves on cross-validation. And specifically the asymptotics for the indirect crossvalidatory bandwidths with K being the Epanechnikov kernel and $L = K_{2r}$, are given below for $r = 1, 2, 8$ and $r \to \infty$. Here $r = 1$ is classical cross-validation (CV) using the

Epanechnikov kernel, and $r \to \infty$ is indirect cross-validation with the Gaussian kernel as the indirect kernel (ICV_G).

$$\begin{aligned}
\sigma^2_{CV} &= C_{f,K}\left\{4R(K)\frac{\mathrm{V}(f'')}{R(f'')R(f)} + 7.42\right\} \\
\sigma^2_{ICV_2} &= C_{f,K}\left\{4R(K)\frac{\mathrm{V}(f'')}{R(f'')R(f)} + 4.71\right\} \\
\sigma^2_{ICV_8} &= C_{f,K}\left\{4R(K)\frac{\mathrm{V}(f'')}{R(f'')R(f)} + 3.72\right\} \\
\sigma^2_{ICV_G} &= C_{f,K}\left\{4R(K)\frac{\mathrm{V}(f'')}{R(f'')R(f)} + 3.48\right\}.
\end{aligned}$$

The first improvement of going from standard cross-validation to having an indirect kernel of one more power is the most important one. The crucial component of the asymptotic theory is decreasing from 7.42 to 4.71. This is sufficiently substantial to consider this simple adjustment of classical cross-validation to solving a good and important part of the problem with the volatility of the cross-validation estimator. However, indirect cross-validation can do better. Going to the Gaussian limit brings the crucial constant down to 3.48! This is quite low and approaching the do-validation constant of 2.19 found in Mammen et al. (2011). It turns out that 3.48 is still so big that 2.19 is a major improvement in theory and practice. Do-validation does better than indirect cross-validation in theory and practice, even when the latter is based on the optimal Gaussian kernel.

2.1 Simulation experiments about indirect cross-validation

The purpose of this section is to study the performance of the indirect cross-validation method with respect to standard cross-validation and the optimal ISE bandwidth $\widehat{h}_{ISE}$ defined in (3). We consider in the study three possible indirect crossvalidatory bandwidths: $\widehat{h}_{ICV_2}$, $\widehat{h}_{ICV_8}$ and $\widehat{h}_{ICV_G}$. These arise by using the Epanechnikov kernel as the kernel K, and as kernel L, the higher power polynomial kernel, K_{2r} defined in (5), for $r = 2, 8$, and K_∞ that is the Gaussian kernel.

We consider the same data generating processes as Mammen et al. (2011). We simulate six designs defined by the six densities plotted in Figure 2 and defined as follows:

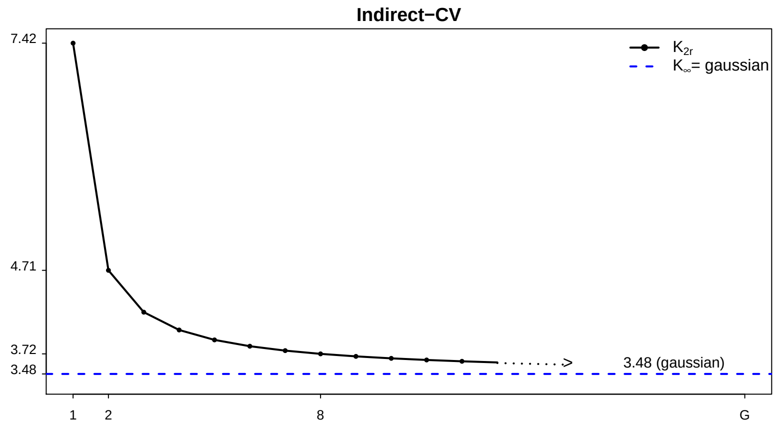

Figure 1: Asymptotic variance reduction for indirect cross-validation with kernels K_{2r} for $r = 1, 2, \ldots \infty$. The limit kernel is the Gaussian plotted using a discontinuous line.

1. a simple normal distribution, $N(0.5, 0.2^2)$,

2. a bimodal mixture of two normals which were $N(0.35, 0.1^2)$ and $N(0.65, 0.1^2)$,

3. a mixture of three normals, namely $N(0.25, 0.075^2)$, $N(0.5, 0.075^2)$ and $N(0.75, 0.075^2)$ giving three clear modes,

4. a gamma distribution, $Gamma(a, b)$ with $b = 1.5$, $a = b^2$ applied on $5x$ with $x \in I\!R_+$, i.e.
$$f(x) = 5\frac{b^a}{\Gamma(a)}(5x)^{a-1}e^{-5xb},$$

5. a mixture of two gamma distributions, $Gamma(a_j, b_j)$, $j = 1, 2$ with $a_j = b_j^2$, $b_1 = 1.5$, $b_2 = 3$ applied on $6x$, i.e.
$$f(x) = \frac{6}{2}\sum_{j=1}^{2}\frac{b_j^{a_j}}{\Gamma(a_j)}(6x)^{a_j-1}e^{-6xb_j}$$
giving one mode and a plateau,

6. and a mixture of three gamma distributions, $Gamma(a_j, b_j)$, $j = 1, 2, 3$ with $a_j = b_j^2$, $b_1 = 1.5$, $b_2 = 3$, and $b_3 = 6$ applied on $8x$ giving two bumps and one plateau.

Our set of densities contains density functions with one, two or three modes, some being asymmetric. They all have exponentially falling tails, because otherwise one

	Design 1					Design 2				
	$\widehat{h}_{ISE}$	$\widehat{h}_{CV}$	$\widehat{h}_{ICV_2}$	$\widehat{h}_{ICV_8}$	$\widehat{h}_{ICV_G}$	$\widehat{h}_{ISE}$	$\widehat{h}_{CV}$	$\widehat{h}_{ICV_2}$	$\widehat{h}_{ICV_8}$	$\widehat{h}_{ICV_G}$
					$n = 100$					
m_1	2.328	4.944	4.804	4.583	4.446	3.477	6.313	6.047	5.876	5.809
m_2	1.876	6.185	5.557	5.451	5.256	1.989	5.611	5.089	4.683	4.511
m_3	0.000	4.963	4.087	3.840	3.472	0.000	2.550	1.969	1.969	1.969
m_4	0.000	-1.839	-1.044	-0.458	-0.230	0.000	0.587	1.049	1.532	1.746
m_5	0.000	0.714	0.724	0.724	0.722	0.000	0.943	0.948	0.950	1.000
					$n = 200$					
m_1	1.417	2.573	2.481	2.359	2.288	2.307	3.816	3.700	3.495	3.451
m_2	1.098	2.747	2.609	2.435	2.213	1.372	3.376	3.239	2.577	2.533
m_3	0.000	3.161	2.545	2.133	1.989	0.000	2.174	1.821	1.487	1.477
m_4	0.000	-0.718	-0.438	0.024	0.192	0.000	-0.087	0.240	0.593	0.769
m_5	0.000	0.687	0.690	0.687	0.687	0.000	0.723	0.751	0.733	0.737
					$n = 500$					
m_1	0.731	1.221	1.175	1.129	1.108	1.208	1.780	1.756	1.695	1.674
m_2	0.465	1.078	1.027	0.913	0.867	0.648	1.237	1.245	1.147	1.122
m_3	0.000	2.615	2.214	1.935	1.818	0.000	1.296	1.218	1.076	0.997
m_4	0.000	-0.805	-0.417	-0.193	-0.104	0.000	-0.195	0.008	0.193	0.285
m_5	0.000	0.666	0.666	0.651	0.656	0.000	0.632	0.629	0.632	0.634
					$n = 1000$					
m_1	0.439	0.719	0.712	0.675	0.664	0.732	1.049	1.006	0.987	0.976
m_2	0.277	0.699	0.699	0.622	0.606	0.377	0.722	0.624	0.609	0.599
m_3	0.000	2.190	2.161	1.741	1.688	0.000	1.227	1.071	0.914	0.857
m_4	0.000	-0.596	-0.434	-0.236	-0.155	0.000	-0.201	-0.074	0.051	0.119
m_5	0.000	0.667	0.667	0.643	0.632	0.000	0.586	0.560	0.554	0.554

Table 1: *Simulation results about the indirect cross-validation method with designs 1 and 2. We compare the standard cross-validation,* $\widehat{h}_{CV}$*, with three indirect versions* $\widehat{h}_{ICV_2}$*,* $\widehat{h}_{ICV_8}$ *and* $\widehat{h}_{ICV_G}$ *for kernels* K_{2r} *with* $r = 2, 8, \infty$*. As a benchmark we report the results for the unfeasible ISE optimal bandwidth,* $\widehat{h}_{ISE}$*. All the numbers have been multiplied by 100.*

has to work with boundary correcting kernels. The main mass is always in $[0, 1]$. For the purposes of this paper we use five measures to summarize the stochastic

	Design 3					Design 4				
	$\widehat{h}_{ISE}$	$\widehat{h}_{CV}$	$\widehat{h}_{ICV_2}$	$\widehat{h}_{ICV_8}$	$\widehat{h}_{ICV_G}$	$\widehat{h}_{ISE}$	$\widehat{h}_{CV}$	$\widehat{h}_{ICV_2}$	$\widehat{h}_{ICV_8}$	$\widehat{h}_{ICV_G}$
						$n = 100$				
m_1	4.448	7.232	7.061	6.905	6.951	4.842	7.918	7.818	7.636	7.643
m_2	2.231	6.392	6.141	5.326	5.247	2.644	6.698	6.842	6.400	6.440
m_3	0.000	1.766	1.515	1.461	1.526	0.000	1.595	1.460	1.357	1.328
m_4	0.000	0.629	1.008	1.423	1.705	0.000	-0.256	1.146	0.569	0.742
m_5	0.000	0.824	0.883	0.957	1.060	0.000	0.822	0.842	0.844	0.869
						$n = 200$				
m_1	2.830	4.216	4.034	3.872	3.864	3.100	4.643	4.521	4.453	4.405
m_2	1.343	3.043	2.788	2.310	2.299	1.657	3.645	3.299	3.391	3.265
m_3	0.000	1.244	1.016	0.947	0.932	0.000	1.396	1.228	1.106	1.118
m_4	0.000	0.086	0.291	0.593	0.707	0.000	-0.313	-0.042	0.233	0.360
m_5	0.000	0.626	0.649	0.626	0.648	0.000	0.758	0.765	0.765	0.778
						$n = 500$				
m_1	1.540	2.006	1.955	1.908	1.889	1.687	2.338	2.270	2.193	2.164
m_2	0.685	1.053	0.998	0.994	0.963	0.767	1.576	1.516	1.333	1.272
m_3	0.000	0.859	0.812	0.673	0.640	0.000	0.924	0.826	0.721	0.717
m_4	0.000	-0.074	0.033	0.187	0.271	0.000	-0.436	-0.205	-0.031	0.073
m_5	0.000	0.562	0.532	0.532	0.532	0.000	0.625	0.611	0.587	0.587
						$n = 1000$				
m_1	0.943	1.166	1.135	1.112	1.109	1.060	1.341	1.317	1.281	1.270
m_2	0.405	0.620	0.590	0.533	0.536	0.491	0.725	0.706	0.645	0.632
m_3	0.000	0.683	0.553	0.457	0.444	0.000	0.784	0.639	0.549	0.504
m_4	0.000	-0.118	0.003	0.088	0.139	0.000	-0.287	-0.132	0.030	0.114
m_5	0.000	0.446	0.467	0.428	0.438	0.000	0.564	0.538	0.528	0.528

Table 2: *Simulation results about the indirect cross-validation method with designs 3 and 4. We compare the standard cross-validation, $\widehat{h}_{CV}$, with three indirect versions $\widehat{h}_{ICV_2}$, $\widehat{h}_{ICV_8}$ and $\widehat{h}_{ICV_G}$ for kernels K_{2r} with $r = 2, 8, \infty$. As a benchmark we report the results for the unfeasible ISE optimal bandwidth, $\widehat{h}_{ISE}$. All the numbers have been multiplied by 100.*

performance of any bandwidth selectors $\widehat{h}$:

$$m_1 = \text{mean}(\text{ISE}(\widehat{h})) \tag{6}$$

$$m_2 = \text{std}(\text{ISE}(\widehat{h})) \tag{7}$$

$$m_3 = 90\% quantile \left(|ISE(\hat{h}) - ISE(\widehat{h}_{ISE})| / ISE(\widehat{h}_{ISE}) \right) \tag{8}$$

$$m_4 = \text{mean}(\widehat{h} - \widehat{h}_{ISE}) \tag{9}$$

$$m_5 = 90\% quantile \left(|\hat{h} - \widehat{h}_{ISE}| / \widehat{h}_{ISE} \right) . \tag{10}$$

	Design 5					Design 6				
	$\widehat{h}_{ISE}$	$\widehat{h}_{CV}$	$\widehat{h}_{ICV_2}$	$\widehat{h}_{ICV_8}$	$\widehat{h}_{ICV_G}$	$\widehat{h}_{ISE}$	$\widehat{h}_{CV}$	$\widehat{h}_{ICV_2}$	$\widehat{h}_{ICV_8}$	$\widehat{h}_{ICV_G}$
						$n = 100$				
m_1	3.356	5.575	5.488	5.250	5.208	3.633	5.458	5.279	5.184	5.149
m_2	1.383	6.160	6.065	4.638	4.575	1.617	4.309	3.829	3.245	3.140
m_3	0.000	1.730	1.624	1.458	1.380	0.000	1.332	1.121	1.122	1.122
m_4	0.000	-0.101	0.616	1.270	1.521	0.000	0.749	1.369	1.970	2.279
m_5	0.000	0.920	0.999	0.999	0.999	0.000	0.917	0.950	0.999	0.999
						$n = 200$				
m_1	2.293	3.516	3.400	3.269	3.223	2.387	3.397	3.317	3.220	3.212
m_2	0.864	2.907	2.483	2.309	2.194	0.955	2.378	2.209	2.014	1.949
m_3	0.000	1.551	1.425	1.289	1.248	0.000	1.160	1.040	0.965	0.962
m_4	0.000	-0.370	0.158	0.638	0.833	0.000	0.373	0.672	1.020	1.194
m_5	0.000	0.791	0.819	0.825	0.807	0.000	0.856	0.839	0.839	0.838
						$n = 500$				
m_1	1.287	1.857	1.806	1.758	1.729	1.355	1.823	1.746	1.719	1.700
m_2	0.520	1.329	1.238	1.176	1.081	0.495	1.150	0.889	0.885	0.821
m_3	0.000	1.298	1.104	0.964	0.930	0.000	0.973	0.886	0.872	0.788
m_4	0.000	-0.093	0.143	0.439	0.602	0.000	-0.333	-0.045	0.244	0.399
m_5	0.000	0.760	0.774	0.751	0.762	0.000	0.637	0.621	0.618	0.618
						$n = 1000$				
m_1	0.844	1.147	1.102	1.075	1.067	0.892	1.074	1.054	1.032	1.024
m_2	0.357	0.691	0.611	0.565	0.546	0.304	0.458	0.445	0.393	0.370
m_3	0.000	1.053	0.866	0.802	0.729	0.000	0.500	0.465	0.403	0.441
m_4	0.000	-0.300	-0.109	0.133	0.245	0.000	-0.149	0.049	0.253	0.381
m_5	0.000	0.667	0.600	0.586	0.591	0.000	0.500	0.516	0.520	0.500

Table 3: *Simulation results about the indirect cross-validation method with designs 5 and 6. We compare the standard cross-validation, $\widehat{h}_{CV}$, with three indirect versions $\widehat{h}_{ICV_2}$, $\widehat{h}_{ICV_8}$ and $\widehat{h}_{ICV_G}$ for kernels K_{2r} with $r = 2, 8, \infty$. As a benchmark we report the results for the unfeasible ISE optimal bandwidth, $\widehat{h}_{ISE}$. All the numbers have been multiplied by 100.*

The above measures have been calculated from 500 simulated samples from each density and four samples sizes $n = 100, 200, 500$ and 1000. The measures m_1, m_2 and m_4 where also used in the simulations by the former paper by Mammen et al. (2011). We have included measures m_3 and m_5 which are informative about the stability of the bandwidth estimates. Tables 1, 2 and 3 show the simulation results. Note that the bias (m_4) is consistently increasing as a function of the power of

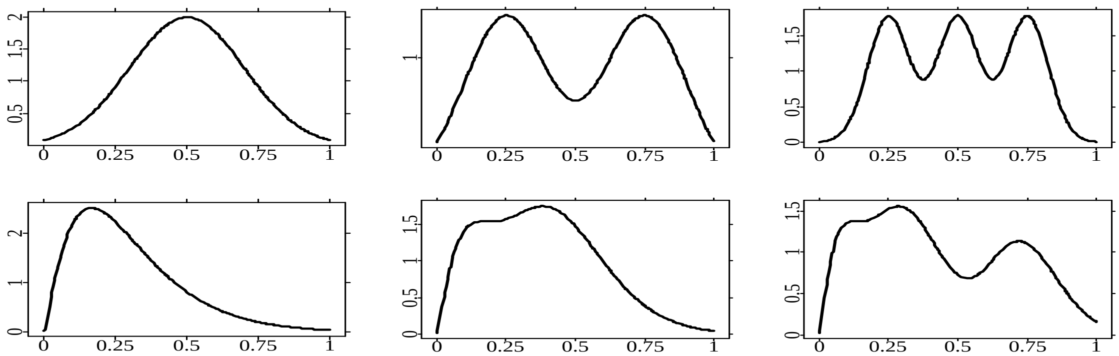

Figure 2: The six data generating densities: Designs 1 to 6 from the upper left to the lower right.

the indirect kernel with the indirect Gaussian kernel having the largest bias. This increase in bias is being more than balanced by a decreasing volatility (m_2) as a function of the power of the indirect kernel. As a result, the overall performance, the averaged integrated squared error performance (m_1), is decreasing as a function of the power of the indirect kernel with the Gaussian indirect kernel performing best of all. These results are very clear for all the designs and sample sizes and the indirectness in cross-validation is indeed working quite well.

3 Indirect do-validation in kernel density estimation

Here we describe the indirect do-validation method and provide theoretical and empirical results in a similar way to that for indirect cross-validation above. We conclude that indirect do-validation improves consistently theoretically when the power of the indirect kernel increases. The relative improvements parallel those we saw for indirect cross-validation. However, it does not seem like the practical improvements follow the theoretical improvements for indirect do-validation. The original conclusion of Mammen et al. (2011) seems to be valid also here: “when the theoretical properties are so good as in do-validation, it is the practical implementation at hand that counts, not further theoretical improvements”. Going all the way to the limiting Gaussian kernel is not of practical relevance for indirect do-validation.

In our derivation of the methodology, we follow Mammen et al. (2011) that first consider a class of bandwidth selectors that are constructed as weighted averages of cross-validation bandwidths. This class of bandwidth selectors contains the classical cross-validation bandwidth selector as one example with $J = 1$ and $L_1(u) = K(u)\mathbf{1}(u \neq 0)$. And it also contains the do-validation method, which combines left and right one-sided cross-validation using the local linear kernel density estimator (Jones, 1993; and Cheng 1997a, 1997b). In fact the method cannot work on local constant density estimation because of its inferior rate of convergence when applying to asymmetric kernels. For a kernel density estimator $\widehat{f}_{h,M}$ with kernel M the local linear kernel density estimator can be defined as kernel density estimator $\widehat{f}_{h,M^*}$ with "equivalent kernel" M^* given by

$$M^*(u) = \frac{\mu_2(M) - \mu_1(M)u}{\mu_0(M)\mu_2(M) - \mu_1^2(M)} M(u).$$

In one-sided cross-validation the basic kernel $M(u)$ is chosen as $2K(u)\mathbf{1}_{(-\infty,0)}$ (left one-sided cross-validation) or $2K(u)\mathbf{1}_{(0,\infty)}$ (right one-sided cross-validation). This results in the left and right one-sided equivalent kernels, K^*_{left} and K^*_{right}, respectively. The left one-sided cross-validation bandwidth is calculated by

$$\left(\frac{R(K)}{R(K^*_{left})} \frac{\mu_2^2(K^*_{left})}{\mu_2^2(K)} \right)^{1/5} \widehat{h}_{K_{left}},$$

where $\widehat{h}_{K_{left}}$ is the minimizer of the left one-sided cross-validation criterion defined as (1), but involving the local linear density estimator with equivalent kernel K^*_{left}. In exactly the same way we define the right one-sided cross-validation bandwidth, but considering now the kernel K^*_{right}. Finally, the do-validation selector $\widehat{h}_{DO}$ is given by the simple average

$$\widehat{h}_{DO} = \left(\frac{R(K)}{R(K^*_{left})} \frac{\mu_2^2(K^*_{left})}{\mu_2^2(K)} \right)^{1/5} \frac{\widehat{h}_{K_{left}} + \widehat{h}_{K_{right}}}{2}. \tag{11}$$

See Martínez-Miranda et al. (2009) and Mammen et al. (2011) for more details. Left one-sided cross-validation and right one-sided cross-validation are not identical in practice because of differences in the boundary. However, asymptotically they are equivalent. As we will see in our simulations do-validation delivers a good stable compromise. It has the same asymptotic theory as each of the two one-sided alternatives and a better overall finite sample performance.

Theorem 1 in Mammen et al. (2011) provides the asymptotic distribution of $\widehat{h}_{DO} - \widehat{h}_{ISE}$. Under their Assumptions (A1) and (A2) it holds for symmetric kernel K that

$$n^{3/10}(\widehat{h}_{DO} - \widehat{h}_{ISE}) \rightarrow N(0, \sigma^2_{DO}) \qquad \text{in distribution,}$$

where σ^2_{DO} has the form of (4). For K equal being the Epanechnikov kernel the asymptotic variance is given by

$$\sigma^2_{\mathrm{DO}} \quad = \quad C_{f,K}\left\{4R(K)\frac{\mathrm{V}(f'')}{R(f'')R(f)} + 2.19\right\}.$$

This can be compared with the asymptotic variance of the plug-in bandwidth which is equal to

$$\sigma^2_{PI} \quad = \quad C_{f,K}\left\{4R(K)\frac{\mathrm{V}(f'')}{R(f'')R(f)} + \ 0.72\right\}.$$

The second term (the only one which differs among bandwidth selectors) was also calculated for the quartic kernel, which is the kernel K_{2r} with $r = 2$. The calculation as above gave the value 1.89 and 0.83 instead of 2.19 and 0.72 (see Mammen et al. 2011). The immediate lesson learned from comparing the asymptotic theory of do-validation of the two kernels considered above is the following: the second term is bigger for plug-in estimator for the quartic kernel than for the Epanechnikov estimator. However, for cross-validation and do-validation it is the exact opposite, the second term is smaller for the quartic kernel than for the Epanechnikov estimator. Therefore, relatively speaking the validation approaches do better for the higher power kernel K_{2r} with $r = 2$, than for the lower power kernel K_{2r}, with $r = 1$ (the Epanechnikov kernel). One could argue that validation does better for the higher power kernel than for the lower power kernel. However, lets further consider the case that we are really interested in the optimal bandwidth for the lower power kernel and we really want to use a validation approach to select that bandwidth, see Mammen et al. (2011) for practical arguments for using validation instead of plug-in. Then it seems intuitively appealing to carry that validation out at the kernel with a high power to select the validated bandwidth for that higher power kernel and then adjusting this bandwidth to the lower power kernel by multiplying by the kernel constant $d_{K_{2r}}^{-1}$ defined in (2). And this is what we call hereafter indirect do-validation to distinguish it from do-validation where the indirect kernel is a one-sided version of K. The formal definition of the indirect do-validation bandwidth $\widehat{h}_{IDO_r}$ with kernels K and K_{2r} is therefore

$$\widehat{h}_{IDO_r} = d_{K_{2r}}^{-1} \widehat{h}_{DO,r}$$

where $\widehat{h}_{DO,r}$ is the do-validation bandwidth defined in (11) but calculated with $K = K_{2r}$ $(r = 1, 2, \ldots)$.

By simple calculations we can write $\widehat{h}_{IDO_r}$ as

$$\widehat{h}_{IDO_r} = \left(\frac{R(K)}{\mu_2^2(K)} \frac{\mu_2^2(K^*_{2r,left})}{R(K^*_{2r,left})} \right)^{1/5} \left(\frac{\widehat{h}_{K_{2r,left}} + \widehat{h}_{K_{2r,right}}}{2} \right),$$

where $\widehat{h}_{K_{2r,left}}$ and $\widehat{h}_{K_{2r,right}}$ are the minimizers of the cross-validation criterion (1) involving the local linear density estimators with equivalent kernels $K^*_{2r,left}$ and $K^*_{2r,right}$, respectively. Then, we can use again Theorem 1 in Mammen et al. (2011) to get that

$$n^{3/10}(\widehat{h}_{IDO_r} - \widehat{h}_{ISE}) \to N(0, \sigma^2_{IDO_r}) \qquad \text{in distribution}$$

where $\sigma^2_{IDO_r}$ has again the same form as (4). We get a result that is similar to the findings in our discussion of indirect cross-validation in Section 2. By increasing the power r $(r = 2, 3, 4, \ldots)$ of the indirect kernel we get an incremental reduction in the asymptotic variance factor. Again, for $r \to \infty$ the factor converges to the factor of indirect do-validation with Gaussian kernel. This can be shown as in Section 2. Figure 3 shows the factor as a function of r.

One sees that the trick of indirect do-validation significantly improves on do-validation. Below we provide the resulting asymptotics for the indirect do-validation bandwidths, h_{IDO_r}, with $r = 1, 2, 8$ and the Gaussian kernel.

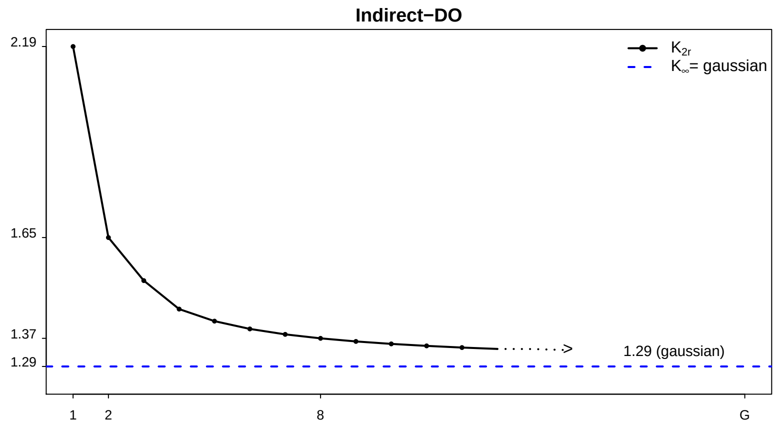

Figure 3: *Asymptotic variance term for indirect do-validation with kernels K_{2r} for $r = 1, 2, \ldots \infty$. The limit kernel is the Gaussian plotted using a discontinuous line.*

$$\begin{aligned}
\sigma^2_{\mathrm{DO}} &= C_{f,K}\left\{4R(K)\frac{\mathrm{V}(f'')}{R(f'')R(f)} + 2.19\right\} \\
\sigma^2_{\mathrm{IDO}_2} &= C_{f,K}\left\{4R(K)\frac{\mathrm{V}(f'')}{R(f'')R(f)} + 1.65\right\} \\
\sigma^2_{\mathrm{IDO}_8} &= C_{f,K}\left\{4R(K)\frac{\mathrm{V}(f'')}{R(f'')R(f)} + 1.37\right\} \\
\sigma^2_{\mathrm{IDO}_\mathrm{G}} &= C_{f,K}\left\{4R(K)\frac{\mathrm{V}(f'')}{R(f'')R(f)} + 1.29\right\}
\end{aligned}$$

3.1 Simulation experiments about indirect do-validation

Here we extend the simulation experiments carried out for indirect cross-validation in Subsection 2.1 with the just defined indirect do-validation method. We evaluate the finite sample performance of highering the powers of the indirect kernel for indirect do-validation, and compare with the former do-validation and the optimal ISE bandwidth ($\widehat{h}_{ISE}$). We consider in the study three possible indirect do-validation bandwidths: $\widehat{h}_{IDO_2}$, $\widehat{h}_{IDO_8}$ and $\widehat{h}_{IDO_G}$, which were defined above.

Tables 4, 5 and 6 show the simulation results. As we saw for indirect classical cross-validation, the finite sample bias (m_4) is consistently increasing when highering the power of the indirect kernel. However, this increase in bias is offset by a decrease in volatility (m_2). This is consistently over sample size and design and follow the results we saw in the previous section for classical cross-validation. However, when

	Design 1					Design 2				
	$\widehat{h}_{ISE}$	$\widehat{h}_{DO}$	$\widehat{h}_{IDO_2}$	$\widehat{h}_{IDO_8}$	$\widehat{h}_{IDO_G}$	$\widehat{h}_{ISE}$	$\widehat{h}_{DO}$	$\widehat{h}_{IDO_2}$	$\widehat{h}_{IDO_8}$	$\widehat{h}_{IDO_G}$
					$n = 100$					
m_1	2.328	3.052	3.038	2.999	2.989	3.477	4.949	5.141	5.504	5.723
m_2	1.876	2.204	2.211	2.172	2.143	1.989	2.642	2.703	2.761	2.757
m_3	0.000	1.324	1.058	1.058	1.060	0.000	1.277	1.385	1.646	1.832
m_4	0.000	1.902	2.290	2.745	3.041	0.000	3.389	4.080	5.171	5.808
m_5	0.000	0.583	0.603	0.616	0.633	0.000	0.914	0.999	1.139	1.204
					$n = 200$					
m_1	1.417	1.803	1.788	1.776	1.775	2.307	2.930	2.925	3.011	3.108
m_2	1.098	1.402	1.373	1.341	1.313	1.372	1.663	1.651	1.668	1.693
m_3	0.000	0.900	0.833	0.851	0.880	0.000	0.748	0.755	0.893	1.029
m_4	0.000	1.116	1.414	1.760	2.022	0.000	1.607	1.865	2.421	2.859
m_5	0.000	0.516	0.532	0.563	0.581	0.000	0.632	0.667	0.750	0.826
					$n = 500$					
m_1	0.731	0.903	0.889	0.878	0.876	1.208	1.439	1.439	1.442	1.458
m_2	0.465	0.559	0.553	0.537	0.532	0.648	0.775	0.773	0.771	0.777
m_3	0.000	0.750	0.690	0.667	0.688	0.000	0.526	0.543	0.568	0.601
m_4	0.000	0.418	0.618	0.836	0.990	0.000	0.679	0.832	1.058	1.258
m_5	0.000	0.464	0.470	0.483	0.500	0.000	0.500	0.524	0.552	0.601
					$n = 1000$					
m_1	0.439	0.525	0.519	0.514	0.513	0.732	0.846	0.841	0.839	0.842
m_2	0.277	0.320	0.316	0.313	0.312	0.377	0.426	0.425	0.425	0.429
m_3	0.000	0.615	0.535	0.491	0.480	0.000	0.459	0.410	0.377	0.410
m_4	0.000	0.297	0.438	0.569	0.659	0.000	0.345	0.423	0.564	0.681
m_5	0.000	0.434	0.432	0.449	0.464	0.000	0.421	0.428	0.448	0.471

Table 4: *Simulation results about the indirect do-validation method with designs 1 and 2. We compare the original do-validated bandwidth, $\widehat{h}_{DO}$, with three indirect versions $\widehat{h}_{IDO_2}$, $\widehat{h}_{IDO_8}$ and $\widehat{h}_{IDO_G}$ for kernels K_{2r} with $r = 2, 8, \infty$. All the numbers have been multiplied by 100.*

it comes to the overall average integrated squared error performance the impression is less clear. Sometimes increasing the power of the indirect kernel improves results, sometimes it does not. Overall, the indirect do-validation methods perform more or less the same. Therefore, for do-validation the decrease in volatility (m_2) and the increase (m_4) seem to be effects of similar size overall. So, the estimators have similar averaged ISE behavior, but they are quite different, when it comes to their

	Design 3					Design 4				
	$\widehat{h}_{ISE}$	$\widehat{h}_{DO}$	$\widehat{h}_{IDO_2}$	$\widehat{h}_{IDO_8}$	$\widehat{h}_{IDO_G}$	$\widehat{h}_{ISE}$	$\widehat{h}_{DO}$	$\widehat{h}_{IDO_2}$	$\widehat{h}_{IDO_8}$	$\widehat{h}_{IDO_G}$
					$n = 100$					
m_1	4.448	11.283	11.597	11.532	11.328	4.842	6.462	6.483	6.536	6.601
m_2	2.231	3.885	3.774	3.904	3.960	2.644	3.246	3.209	3.208	3.216
m_3	0.000	4.913	4.996	4.913	4.782	0.000	0.943	0.941	1.021	1.027
m_4	0.000	10.198	10.687	10.793	10.661	0.000	3.352	3.619	4.019	4.300
m_5	0.000	1.943	1.943	1.943	1.943	0.000	0.894	0.944	1.000	1.027
					$n = 200$					
m_1	2.830	3.799	3.997	4.207	4.391	3.100	3.940	3.955	3.956	3.984
m_2	1.343	2.288	2.560	2.624	2.630	1.657	2.032	2.018	1.983	1.980
m_3	0.000	0.811	0.941	1.256	1.368	0.000	0.774	0.797	0.830	0.871
m_4	0.000	1.895	2.307	2.899	3.345	0.000	2.147	2.371	2.670	2.904
m_5	0.000	0.734	0.896	1.085	1.159	0.000	0.794	0.853	0.912	0.922
					$n = 500$					
m_1	1.540	1.757	1.751	1.767	1.798	1.687	1.967	1.956	1.961	1.974
m_2	0.685	0.815	0.806	0.808	0.829	0.767	0.882	0.877	0.873	0.870
m_3	0.000	0.397	0.390	0.417	0.458	0.000	0.491	0.456	0.511	0.548
m_4	0.000	0.545	0.627	0.839	1.036	0.000	0.973	1.108	1.368	1.546
m_5	0.000	0.438	0.440	0.498	0.531	0.000	0.587	0.587	0.617	0.632
					$n = 1000$					
m_1	0.943	1.044	1.039	1.039	1.045	1.060	1.174	1.169	1.175	1.183
m_2	0.405	0.449	0.450	0.454	0.462	0.491	0.534	0.523	0.518	0.517
m_3	0.000	0.300	0.274	0.270	0.281	0.000	0.322	0.315	0.326	0.345
m_4	0.000	0.206	0.279	0.408	0.517	0.000	0.544	0.662	0.870	1.008
m_5	0.000	0.368	0.368	0.400	0.435	0.000	0.470	0.470	0.498	0.502

Table 5: *Simulation results about the indirect do-validation method with designs 3 and 4. We compare the original do-validated bandwidth, $\widehat{h}_{DO}$, with three indirect versions $\widehat{h}_{IDO_2}$, $\widehat{h}_{IDO_8}$ and $\widehat{h}_{IDO_G}$ for kernels K_{2r} with $r = 2, 8, \infty$. All the numbers have been multiplied by 100.*

bias/variance trade off.

4 Concluding remarks

This paper is on indirect cross-validation. The term indirect cross-validation originates from one particular application of it by Savchuk et al. (2010). The do-validation version of indirect cross-validation introduced in Mammen et al. (2011)

	Design 5					Design 6				
	$\widehat{h}_{ISE}$	$\widehat{h}_{DO}$	$\widehat{h}_{IDO_2}$	$\widehat{h}_{IDO_8}$	$\widehat{h}_{IDO_G}$	$\widehat{h}_{ISE}$	$\widehat{h}_{DO}$	$\widehat{h}_{IDO_2}$	$\widehat{h}_{IDO_8}$	$\widehat{h}_{IDO_G}$
$n = 100$										
m_1	3.356	4.437	4.509	4.533	4.539	3.633	4.972	5.036	5.136	5.211
m_2	1.383	1.566	1.534	1.476	1.446	1.617	1.911	1.884	1.850	1.817
m_3	0.000	1.061	1.137	1.185	1.185	0.000	1.000	1.031	1.078	1.157
m_4	0.000	5.434	6.041	6.557	6.721	0.000	5.794	6.305	6.913	7.317
m_5	0.000	0.999	0.999	1.021	1.042	0.000	1.042	1.084	1.131	1.174
$n = 200$										
m_1	2.293	3.008	3.036	3.054	3.063	2.387	3.206	3.250	3.310	3.362
m_2	0.864	1.002	0.984	0.918	0.896	0.955	1.242	1.270	1.272	1.273
m_3	0.000	1.029	1.050	1.048	1.065	0.000	0.971	1.006	1.052	1.086
m_4	0.000	4.331	4.707	5.149	5.378	0.000	4.163	4.508	5.018	5.386
m_5	0.000	0.923	0.923	0.977	0.999	0.000	0.975	0.999	1.067	1.090
$n = 500$										
m_1	1.287	1.668	1.678	1.695	1.710	1.355	1.694	1.700	1.718	1.738
m_2	0.520	0.550	0.537	0.520	0.512	0.495	0.621	0.617	0.604	0.596
m_3	0.000	0.911	0.958	0.990	1.027	0.000	0.712	0.714	0.709	0.750
m_4	0.000	3.211	3.449	3.820	4.046	0.000	2.430	2.624	2.935	3.148
m_5	0.000	0.875	0.928	0.951	0.976	0.000	0.751	0.751	0.786	0.800
$n = 1000$										
m_1	0.844	1.016	1.023	1.036	1.050	0.892	1.030	1.031	1.043	1.056
m_2	0.357	0.396	0.393	0.381	0.374	0.304	0.334	0.327	0.318	0.315
m_3	0.000	0.612	0.632	0.717	0.771	0.000	0.509	0.480	0.492	0.543
m_4	0.000	1.868	2.064	2.387	2.605	0.000	1.448	1.580	1.878	2.077
m_5	0.000	0.718	0.728	0.775	0.799	0.000	0.600	0.595	0.638	0.684

Table 6: *Simulation results about the indirect do-validation method with designs 5 and 6. We compare the original do-validated bandwidth, $\widehat{h}_{DO}$, with three indirect versions $\widehat{h}_{IDO_2}$, $\widehat{h}_{IDO_8}$ and $\widehat{h}_{IDO_G}$ for kernels K_{2r} with $r = 2, 8, \infty$. All the numbers have been multiplied by 100.*

proved superior to earlier versions of indirect cross-validation in practice. See the empirical work in Savchuk et al. (2008), where the original indirect cross-validated bandwidth selector had inferior performance to the celebrated plug-in type bandwidth selector of Sheather and Jones (1991). However, the numerical work in Mammen et al. (2011) indicated that do-validation performs better in practice than a plug-in bandwidth selector. We also considered plug-in bandwidths and various combinations of plug-in, cross-validated and do-validated bandwidths. Neither of these attempts performed as well as the do-validated bandwidth. It would be interesting to see whether indirect cross-validation and do-validation would also be useful to improve other variants of the kernel density estimation problem, such as the problem considered by Gavriliadis and Athanassoulis (2012), Park (2013), Eidous (2012) or Martínez-Miranda et al. (2013).

Acknowledgements

This research has been partially supported by the Spanish “Ministerio de Ciencia e Innovación”, grant MTM2008-03010, the Swiss National Science Foundation, project 100018-140295, and the European Commission under the Marie Curie Intra-European Fellowship FP7-PEOPLE-2011-IEF project number 302600.

References

Buch-Kromann T., Nielsen, J.P., 2012. Multivariate density estimation using dimension reducing information and tail flattening transformations for truncated or censored data, *Annals of the Institute of Mathematical Statistics*, **64** (1), 167–192.

Cheng, M.Y., 1997a. Boundary-aware estimators of integrated squared density derivatives. *Journal of the Royal Statistical Society Ser. B*, **50**, 191–203.

Cheng, M.Y., 1997b. A bandwidth selector for local linear density estimators. *The Annals of Statistics*, **25**, 1001–1013.

Eidous, O.M. 2012. A new kernel estimator for abundance using line transect sampling without the shoulder condition. *Journal of the Korean Statistical Society*, **41**(2), 267-–275.

Gámiz-Pérez, M.L., Martínez-Miranda, M.D., Nielsen, J.P., 2013a. Smoothing survival densities in practice. *Computational Statistics and Data Analysis*, **58**, 368-–382.

Gámiz-Pérez, M.L., Janys, L., Martínez-Miranda, M.D., Nielsen, J.P., 2013b. Bandwidth selection in marker dependent kernel hazard estimation. *Computational Statistics and Data Analysis*, **68**, 155-–169.

Gámiz-Pérez, M.L., Mammen, E., Martínez-Miranda, M.D., Nielsen, J.P., 2013c. Do-validating local linear hazards. *Submitted preprint.*

Gavriliadis, P.N. and Athanassoulis, G.A., 2012, The truncated Stieltjes moment problem solved by using kernel density functions. *Journal of Computational and Applied Mathematics*, **236**(17), 4193–4213.

González-Manteiga W., Lombardía-Cortiña, M.J., Martínez-Miranda, M.D., Sperlich, S., 2013. Kernel Smoothers and Bootstrapping for Semiparametric Mixed Effects Models, *Journal of Multivariate Analysis*, **114**, 288–302.

Hart, J.D., Lee, C.-L., 2005. Robustness of one-sided cross-validation to autocorrelation. *Journal of Multivariate Statistics*, **92**, 77–96.

Hart, J.D., Yi, S., 1998. One-Sided Cross-Validation. *Journal of the American Statistical Association*, **93**, 620–631.

Jones, M.C., 1993. Simple boundary correction in kernel density estimation. *Statistics and Computing*, **3**, 135–146.

Lee, Y.K., Mammen, E., Park, B.U., 2010. Backfitting and smooth backfitting for additive quantile models. *Annals of Statistics*, **38** (5), 2857–2883.

Lee, Y.K., Mammen, E., Park, B.U., 2012a. Flexible generalized varying coefficient regression models. *Annals of Statistics*, **40** (3), 1906–1933.

Lee, Y.K., Mammen, E., Park, B.U., 2012b. Projection-type estimation for varying coefficient regression models. *Bernoulli*, **18** (1), 177–205.

Mammen, E., Martínez-Miranda, M.D., Nielsen, J.P., Sperlich, S., 2011. Do-validation for kernel density estimation, *Journal of the American Statistical Association*, **106**, 651–660.

Martínez-Miranda, M.D., Nielsen, J.P., Sperlich, S., 2009. One sided crossvalidation for density estimation with an application to operational risk. *In "Operational Risk Towards Basel III: Best Practices and Issues in Modelling. Management and Regulation, ed. G.N. Gregoriou; John Wiley and Sons, Hoboken, New Jersey.*

Martínez-Miranda, M.D., Nielsen, J.P., Sperlich, S., Verrall, R., 2013. Continuous Chain Ladder: Reformulating and generalizing a classical insurance problem. *Expert Systems with Applications*, **40** (14) 5588–5603.

Oliveira, M., Crujeiras, R.M., Rodríguez-Casal, A., 2012. A plug-in rule for bandwidth selection in circular density estimation. *Computational Statistics and Data Analysis*, **56** (12), 3898–3908.

Park, B.U., Marron, J.S., 1990. Comparison of Data-Driven Bandwidth Selectors. *Journal of the American Statistical Association*. **85**, 66–72.

Park, H.S., 2013. Comparison of relative efficiency of kernel density estimator with the exponential map. *Journal of the Korean Statistical Society*, **42**(2), 267-–275.

Savchuk, O.Y., Hart, J.D., Sheather S.J., 2008. An empirical study of indirect cross-validation. *IMS Lecture Notes - Festschrift for Tom Hettmansperger.*

Savchuk, O.Y., Hart, J.D., Sheather S.J., 2010. Indirect crossvalidation for Density Estimation. *Journal of the American Statistical Association*, **105**, 415–423.

Sheather, S.J., Jones, M.C., 1991. A reliable data-based bandwidth selection method for kernel density estimation. *Journal of the Royal Statistical Society*, Ser. B, **53**, 683–690.

Soni, P., Dewan, I., Jain, K., 2012. Nonparametric estimation of quantile density function. *Computational Statistics and Data Analysis*, **56** (12), 3876—3886.

Spreeuw, J., Nielsen, J.P., Jarner, S.F., 2013. A visual test of mixed hazard models, *SORT-Statistics and Operations Research Transactions*, in press.