

ESTIMASI INDIKATOR KEMISKINAN TINGKAT KECAMATAN MENGUNAKAN REGRESI KEKAR M-KUANTIL

Nia Aprillyana

Badan Pusat Statistik

nia_aprillyana@bps.go.id

Diterima: Maret 2019; Direvisi: April 2019; Disetujui: Mei 2019

Abstract. *The poverty indicators is one of the strategic figures in Indonesia to determine country development policies. The poverty indicators is divided into three, namely the Headcount Ratio (HCR), Poverty Gap (PG), and Poverty Severity (PS). Through the National Socio-Economic Survey (Susenas), Statistics Indonesia (BPS) is only able to estimate poverty indicators until the district level. On the other hand, sub-district poverty indicators are needed by regional governments to determine the regional policies. Some branch of BPS conduct an additional survey (namely Suseda) to increase the number of samples so they could estimate until the sub-district level. However, the increasing of the number of samples could not be conduct by all branch of BPS because of the constrained costs. To solve the problem, we need a model-based small area estimation technique which only uses original samples of Susenas. The small area estimation technique used the M-quantile regression model which robust to the outliers. This model used the influence function and the quantile value in weighting the residual to produce an estimate of the regression coefficient. In the last part, this study calculated the comparison of accuracy between the increasing of the number of samples and the use of models in producing poverty indicators at the sub-district level. To evaluate accuracy, we use RRMSE for each sub-district poverty indicators estimation.*

Keywords: *influence function, m-quantile, poverty indicator, small area estimation.*

Abstraksi. *Indikator kemiskinan merupakan salah satu angka strategis di Indonesia untuk menentukan kebijakan pembangunan. Indikator kemiskinan terbagi menjadi tiga yaitu Angka Kemiskinan (P_0), Kedalaman Kemiskinan (P_1), dan Keparahan Kemiskinan (P_2). Badan Pusat Statistik (BPS) melalui sampel Survei Sosial Ekonomi Nasional (Susenas) hanya mampu mengestimasi angka indikator kemiskinan sampai tingkat kabupaten. Disisi lain, angka kemiskinan tingkat kecamatan sangat dibutuhkan pemerintah daerah untuk menentukan kebijakan daerah. BPS kabupaten beberapa mengadakan survei tambahan (Suseda) untuk menambah sampel sehingga bisa mengestimasi sampai tingkat kecamatan. Akan tetapi, penambahan sampel terkadang tidak bisa dilakukan semua kabupaten karena terkendala biaya. Untuk menangani masalah tersebut, dibutuhkan teknik estimasi area kecil berbasis model yang hanya menggunakan sampel yang ada di Susenas saja tanpa penambahan sampel untuk memperoleh angka indikator kemiskinan tingkat kecamatan. Penelitian ini ingin mengetahui perbandingan akurasi antara adanya penambahan sampel dengan pemakaian model dalam menghasilkan angka indikator kemiskinan tingkat kecamatan. Teknik estimasi area kecil yang digunakan adalah model regresi M-kuantil yang kekar terhadap outlier. Model ini memakai fungsi pengaruh dan nilai kuantil dalam memboboti residual untuk menghasilkan estimasi koefisien regresi. Untuk mengevaluasi akurasi, digunakan RRMSE untuk setiap estimasi indikator kemiskinan tingkat kecamatan.*

Kata kunci: *fungsi pengaruh, m-kuantil, indikator kemiskinan, estimasi area kecil.*

PENDAHULUAN

Survei nasional memiliki peran penting dalam menghasilkan indikator strategis untuk mendukung pembangunan suatu negara. Tujuan utama dari survei nasional adalah untuk menduga parameter sebuah populasi. Survei nasional hanya menyediakan informasi terbatas dan hanya menduga populasi atau sub populasi besar sebagaimana dinyatakan Molina dan Rao (2010). Survei nasional memiliki varian yang besar dan akurasi yang rendah ketika mengestimasi level area yang lebih kecil (seperti kecamatan atau desa) karena ketidakcukupan sampel.

Untuk meningkatkan akurasi, pelaksana survei akan menambah jumlah sampel atau yang disebut dengan *oversampling*. *Oversampling* memiliki dua manfaat yaitu untuk menyediakan pendugaan yang akurat dan untuk menangani non-respon (Kennickell 2007). Akan tetapi *oversampling* tidak efisien karena membutuhkan biaya yang lebih besar, sehingga dibutuhkan alternatif teknik estimasi lain dengan hanya menggunakan sampel kecil (Giusti, Marchetti, Pratesi, & Salvati 2012) yang disebut dengan estimasi tidak langsung atau estimasi yang berbasis pada model. Rao menyatakan bahwa estimasi tidak langsung dapat diimplementasikan dengan “meminjam informasi” atau memanfaatkan variabel-variabel tambahan (variabel penyerta) yang didapatkan pada area lain dari survei yang sama, dari area yang sama pada survei terdahulu, atau variabel lain yang berhubungan dengan variabel yang menjadi perhatian pada area kecil (Rao 2003).

Salah satu indikator pembangunan strategis yang angkanya hanya tersedia untuk level nasional, provinsi, dan kabupaten/kota adalah indikator kemiskinan. Di Indonesia, indikator

kemiskinan secara resmi dihitung oleh BPS melalui Survei Sosial Ekonomi Nasional (Susenas) untuk estimasi tingkat kabupaten, provinsi, dan nasional (BPS, 2016).

Pemerintah daerah sebagai pelaksana kebijakan tingkat daerah tentu membutuhkan indikator kemiskinan hingga level area kecil (kecamatan dan desa). Indikator kemiskinan area kecil biasa diperoleh dengan penambahan sampel di luar survei utama Susenas, atau disebut dengan Survei Sosial Ekonomi Daerah (Suseda). Salah satu kabupaten yang melaksanakan Suseda sekaligus menjadi objek/wilayah penelitian ini adalah Kabupaten Musi Rawas, Provinsi Sumatera Selatan. Di awal pendahuluan telah dijelaskan bahwa penambahan sampel sulit dilakukan karena sejumlah batasan (biaya dan waktu), sehingga penelitian ini mencari alternatif dalam penghitungan indikator kemiskinan yang berbasis pada model.

Penelitian ini fokus pada model M-kuantil karena karakteristik data pengeluaran perkapita rumah tangga pada objek penelitian terbukti *heavy tailed* atau menceng ke kanan. M-kuantil adalah suatu metode semiparametrik yang berbasis kuantil yang menjamin estimasi yang bersifat kekar dan tidak memerlukan asumsi sebaran teoritis sebagaimana dinyatakan oleh Chambers dan Tzavidis (2006). M-kuantil tidak bergantung pada asumsi distribusi yang kuat dan secara otomatis menyediakan kesimpulan yang kekar terhadap pencilan (Chambers & Tzavidis, 2006; Tzavidis, Marchetti & Chambers, 2010; Marchetti, Tzavidis & Pratesi, 2012). M-kuantil menggunakan M-estimator yang memberikan bobot kecil terhadap pencilan (dan residual) dan mengabaikan asumsi yang harus ada pada Metode Kuadrat Terkecil (MKT). Beberapa M-estimator yang bisa dimanfaatkan dalam model M-

kuantil yaitu Huber dan Hampel. Beberapa studi mengenai M-kuantil lebih sering menggunakan fungsi pengaruh Huber dibandingkan dengan fungsi lainnya sebagaimana yang dilakukan oleh Giusti, Marchetti, Pratesi & Salvati (2012) dan Girinoto (2017).

Di akhir analisis, penelitian ini mengevaluasi akurasi estimasi indikator kemiskinan yang dihasilkan M-kuantil dengan akurasi estimasi indikator kemiskinan yang dihasilkan estimasi langsung pada sampel yang telah ditambahkan (Suseda). Apabila hasilnya sama atau lebih akurat, maka teknik M-kuantil bisa menjadi alternatif dalam estimasi kemiskinan di area kecil. Selain itu, penelitian juga membandingkan dua fungsi pengaruh (Huber dan Hampel) dalam proses pembentukan model. Fungsi mana yang lebih menjamin untuk menghasilkan model yang paling akurat.

METODE PENELITIAN

1. Landasan Teori

Kemiskinan dipandang sebagai ketidakmampuan dari sisi ekonomi untuk memenuhi kebutuhan dasar makanan dan bukan makanan yang diukur dari sisi pengeluaran (BPS 2016). Garis Kemiskinan (GK) merupakan penjumlahan dari Garis Kemiskinan Makanan (GKM) dan Garis Kemiskinan Non Makanan (GKNM). GKM merupakan nilai pengeluaran kebutuhan minimum makanan yang disetarakan dengan 2100 kilo kalori perkapita perhari. Paket komoditi kebutuhan dasar makanan diwakili oleh 52 jenis komoditi (padi-padian, umbi-umbian, ikan, daging, telur dan susu, sayuran, kacang-kacangan, buah-buahan, minyak dan lemak, dll). GKNM adalah kebutuhan minimum untuk perumahan, sandang, pendidikan, dan kesehatan.

Berdasarkan pendekatan kebutuhan dasar, BPS menggunakan tiga indikator kemiskinan, yaitu :

Head Count Ratio (HCR– P_0), yang disebut sebagai persentase penduduk miskin, merupakan persentase penduduk yang berada di bawah garis kemiskinan (GK).

Poverty Gap Index (PG– P_1), yang disebut sebagai indeks kedalaman kemiskinan yang merupakan ukuran rata-rata kesenjangan pengeluaran masing-masing penduduk miskin terhadap garis kemiskinan. Semakin tinggi nilai indeks, semakin jauh rata-rata pengeluaran penduduk dari garis kemiskinan.

Poverty Severity Index (PS– P_2), yang disebut sebagai indeks keparahan kemiskinan yang memberikan gambaran mengenai penyebaran pengeluaran di antara penduduk miskin. Semakin tinggi nilai indeks, semakin tinggi ketimpangan pengeluaran di antara penduduk miskin.

Foster *et al.* (1984) merumuskan suatu ukuran untuk mengukur tingkat kemiskinan untuk area kecil d. Rumus berikut disesuaikan dengan konteks penelitian:

$$F_{\alpha,d} = \frac{1}{N_d} \sum_{j=1}^{N_d} \left(\frac{t-y_{jd}}{t} \right)^{\alpha} I(y_{jd} \leq t) \quad (1)$$

Di mana :

d = kecamatan ($d = 1, 2, \dots, D$)

α = koefisien sensitivitas (0 untuk P_0 , 1 untuk P_1 , 2 untuk P_2)

y_{jd} = pengeluaran per kapita unit j di area d

t = Garis Kemiskinan (GK)

N_d = Banyaknya populasi rumah tangga di kecamatan d

$I(y_{jd} \leq t)$ fungsi indikator suatu unit rumah tangga dibawah GK.

$I(y_{jd} \leq t) = 1$ jika $y_{jd} \leq t$ dan $I(y_{jd} \leq t) = 0$ jika $y_{jd} > t$.

Di tahap awal penelitian, kami menghitung indikator kemiskinan dengan metode estimasi langsung. Estimasi dikatakan langsung apabila estimasi terhadap parameter populasi di suatu area hanya didasarkan pada data sampel yang diperoleh dari area tersebut. Pada metode estimasi langsung, penghitungan estimasi P_0 dan P_1 tiap kecamatan- d hanya menggunakan y_{jd} dari gugus sampel. Gugus sampel dalam penghitungan estimasi langsung merupakan penggabungan sampel Susenas dan Suseda (*oversampling*).

Estimasi langsung pada penghitungan indikator kemiskinan merujuk pada rumus umum FGT, menjadi:

$$\hat{P}_{\alpha,d}^{Dir} = \frac{1}{n_d} \sum_{j=1}^{n_d} \left(\frac{t-y_{jd}}{t} \right)^\alpha I(y_{jd} \leq t) \quad (2)$$

Di mana:

d = kecamatan di Kabupaten Musi Rawas ($d = 1, 2, \dots, 14$)

α = koefisien sensitivitas (0 untuk P_0 , 1 untuk P_1)

y_{jd} = pengeluaran perkapita rumah tangga j di kecamatan d

t = Garis Kemiskinan (GK) di kecamatan d

n_d = Banyaknya sampel rumah tangga di kecamatan d

$I(y_{jd} \leq t)$ fungsi indikator suatu unit dibawah GK di kecamatan d

Setelah melakukan estimasi langsung, penghitungan selanjutnya menggunakan model M-kuantil. M-kuantil menggunakan data Susenas 2015 sebagai penyedia variabel respon dalam pemodelan. Selain itu, penghitungan estimasi indikator kemiskinan P_0 , P_1 , dan RRMSE melalui M-kuantil membutuhkan variabel penyerta yang diperoleh dari data Sensus Penduduk 2010 (SP2010) di level rumah tangga. Variabel penyerta dipilih berdasarkan

hubungan dengan variabel respon dan disesuaikan dengan ketersediaan data populasi SP2010.

Model M-kuantil diperoleh dengan mendekomposisi persamaan:

$$\hat{P}_{\alpha,d} = N_d^{-1} (\sum_{j \in s_d} \hat{P}_{\alpha,jd} + \sum_{j \in r_d} \hat{P}_{\alpha,kd}) \quad (3)$$

Di mana:

$\hat{P}_{\alpha,jd}$ adalah estimasi indikator kemiskinan untuk gugus sampel

$\hat{P}_{\alpha,kd}$ adalah estimasi indikator kemiskinan untuk gugus non-sampel

N_d^{-1} = Populasi rumah tangga kecamatan d yang nilainya didekatkan dengan populasi SP2010

r_d adalah gugus rumah tangga non sampel yang nilai indikator kemiskinannya diestimasi melalui model M-kuantil. Untuk mempertimbangkan resiko bias pada estimasi di gugus r_d , model merujuk pada model estimasi distribusi Dunstan dan Chambers (1989)

$$\hat{F}_d(t) = N_d^{-1} \{ \sum_{j \in s_d} I(y_{jd} \leq t) + \sum_{k \in r_d} n_d^{-1} \sum_{j \in s_d} I\{ \hat{y}_{kd} + (y_{jd} - \hat{y}_{jd}) \leq t \} \} \quad (4)$$

Persamaan di atas berbasis pada model linear $Q_{q_{jd}}(\mathbf{X}; \psi) = \mathbf{X}_{jd}^T \beta_\psi(q_{jd})$, regresi M-kuantil order ke- q untuk distribusi bersyarat y_{jd} pada nilai \mathbf{X}_{jd} . Selanjutnya, estimasi indikator kemiskinan pada area- d diperoleh melalui pendekatan rumus sebelumnya dan dikombinasikan dengan rumus umum kemiskinan, sehingga menjadi:

$$\hat{P}_{\alpha,d}(t) = N_d^{-1} \sum_{j \in s_d} \left(\frac{t-y_{jd}}{t} \right)^\alpha I(y_{jd} \leq t) + N_d^{-1} \sum_{k \in r_d} n_d^{-1} \sum_{j \in s_d} \left(\frac{t-(\hat{y}_{kd}+(y_{jd}-\hat{y}_{jd}))}{t} \right)^\alpha I((\hat{y}_{kd} + (y_{jd} - \hat{y}_{jd})) \leq t) \quad (5)$$

$$\text{Untuk } \hat{y}_{jd} = \mathbf{X}_{jd}^T \hat{\beta}_\psi(\hat{\theta}_d) \text{ dan } \hat{y}_{kd} = \mathbf{X}_{kd}^T \hat{\beta}_\psi(\hat{\theta}_d)$$

Di mana:

\hat{y}_{jd} adalah estimasi pengeluaran perkapita rumah tangga pada gugus sampel.

X_{jd}^T adalah variabel penyerta yang diperoleh dari data SP2010 untuk menghitung estimasi pengeluaran perkapita di gugus sampel.

\hat{y}_{kd} adalah estimasi pengeluaran perkapita rumah tangga pada gugus non-sampel.

X_{kd}^T adalah variabel penyerta yang diperoleh dari data SP2010 untuk

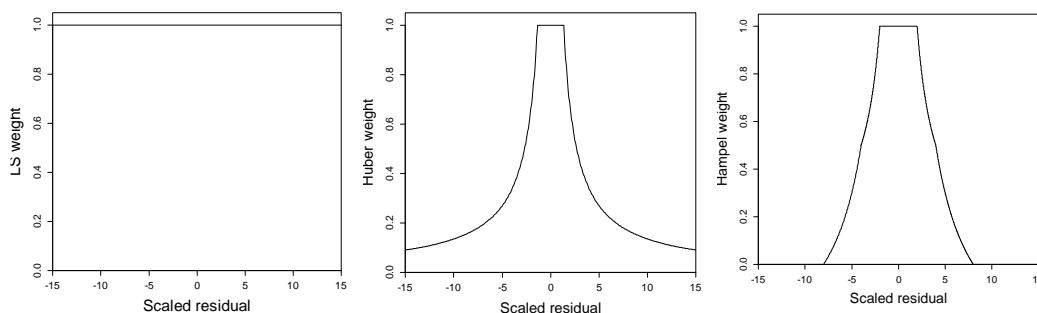
menghitung estimasi pengeluaran perkapita di gugus non-sampel.

$\hat{\beta}_\psi(\hat{\theta}_d)$ adalah estimasi koefisien regresi kecamatan- d yang pada model M-kuantil diperoleh melalui *Iterative Resample Least Square* (IRLS) dengan memasukkan fungsi pengaruh (ψ) yang kontinu. Fungsi pengaruh yang dipakai peneliti adalah tipe M-estimator yaitu Huber dan Hampel yang berguna untuk mengontrol nilai skala terhadap residual dan menjamin model bersifat kekar terhadap data pencilan.

Tabel 1.

Nilai *tuning constant*, scaled residual, dan pembobot pada fungsi pengaruh MKT, Huber, dan Hampel

Fungsi pengaruh dan <i>tuning constant</i>	Nilai scaled residual U	Rumus Pembobot $w(u)$
MKT	$-\infty \leq u \leq \infty$	1
Huber $k = 1.345$	jika $ u < k$	1
	jika $ u \geq k$	$\frac{k}{ u }$
Hampel $a = 2$ $b = 4$ $c = 8$	jika $ u < a$	1
	jika $a \leq u < b$	$\frac{a}{ u }$
	jika $b \leq u \leq c$	$a \frac{\frac{c}{ u } - 1}{c - b}$
	lainnya	0



Gambar 1. Fungsi pembobot Huber, Hampel dan MKT

Sumbu x pada Gambar 1 menunjukkan nilai scaled residual, sedangkan sumbu y menunjukkan nilai pembobot dari MKT, Huber, dan Hampel. Fungsi pembobot Hampel memberikan bobot nol terhadap nilai scaled residual tertentu, sedangkan MKT dan Huber tidak pernah memberikan bobot nol pada nilai scaled residual. MKT selalu memberikan bobot satu untuk semua nilai scaled residual. Gambar 1 menunjukkan bahwa fungsi pembobot M-estimator memberikan nilai bobot yang semakin kecil pada bagian ekor sebaran. Hal ini berbeda dengan MKT yang selalu memberi nilai bobot sebesar satu untuk semua amatannya. Artinya, residual yang bernilai besar memiliki pengaruh yang lebih

kecil terhadap pendugaan kekar dibandingkan jika memakai MKT.

2. Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang berasal dari Susenas Maret 2015, Suseda September 2015, dan Sensus Penduduk 2010 (SP2010) di Kabupaten Musi Rawas Provinsi Sumatera Selatan. Suseda digunakan untuk menambah jumlah sampel dan sekaligus membentuk model M-kuantil, sedangkan SP2010 digunakan sebagai penyedia variabel penyerta yang digunakan untuk mengestimasi nilai pengeluaran perkapita populasi sekaligus mengestimasi indikator kemiskinan.

Tabel 2.
Rincian sampel rumah tangga masing-masing kecamatan

Nama Kecamatan	Susenas	Suseda	Susenas+Suseda
	Tanpa <i>Oversampling</i>		<i>Oversampling</i>
STL Ulu	30	90	120
Selangit	20	90	110
Sumber Harta	30	90	120
Tugumulyo	88	80	168
Purwodadi	30	89	119
Muara Beliti	20	90	110
TP Kepungut	10	90	100
Jayaloka	20	90	110
Suka Karya	30	90	120
Muara Kelingi	30	80	110
BTS Ulu	36	80	116
Tuah Negeri	30	80	110
Muara Lakitan	63	80	143
Megang Sakti	67	80	147
Jumlah	504	1199	1703

Penelitian ini menggunakan tiga variabel penyerta yang diperoleh menggunakan metode *stepwise* dan mempertimbangkan ketersediaan dan kesamaannya dengan variabel yang ada di SP2010, Susenas dan Suseda. Ketiga variabel terpilih tersebut yaitu jumlah anggota rumah tangga (X_1),

lapangan usaha/pekerjaan KRT (X_2), dan status/kedudukan KRT dalam usaha/pekerjaan (X_3). Variabel respon yang diamati pada penelitian ini adalah pengeluaran perkapita rumah tangga per bulan.

Tabel 3.
Daftar variabel yang digunakan beserta penjelasannya

No	Nama Variabel	Kode	Satuan/Rincian	Penjelasan
1.	Pengeluaran per kapita rumah tangga per bulan	Y	Rupiah	Susenas 2015 KP R25
2.	Jumlah anggota rumah tangga	X ₁	Orang	Susenas 2015 Kor Blok II R1, SP 2010 Blok I
3.	Lapangan usaha/pekerjaan kepala rumah tangga	X ₂	1 = pertanian 0 = non pertanian	Susenas 2015 Kor R30, SP 2010 R217
4.	Status dalam pekerjaan	X ₃	1= berusaha sendiri/ dibantu buruh 0= selain berusaha sendiri/dibantu buruh	Susenas 2015 Kor R31, SP 2010 R218

3. Tahapan Analisis

Indikator kemiskinan dengan estimasi langsung mengikuti rumus FGT (1984):

$$\hat{P}_{\alpha,d}^{Dir} = \frac{1}{n_d} \sum_{j=1}^{n_d} \left(\frac{t-y_{jd}}{t} \right)^\alpha I(y_{jd} \leq t) \quad (6)$$

Sedangkan RMSE ($P_{\alpha,d}^{Dir}$) pada estimasi langsung dihitung dengan menggunakan metode bootstrap dengan *resampling* sebanyak 100 kali (B=100). Metode bootstrap menghasilkan $\hat{P}_{\alpha,d}^{*B}$ untuk menghitung RMSE pada estimasi langsung sesuai rumus:

$$\widehat{RMSE}(P_{\alpha,d}^{Dir}) = \sqrt{B^{-1} \sum_{b=1}^B (\hat{P}_{\alpha,d}^{*B} - \hat{P}_{\alpha,d}^{Dir})^2} \quad (7)$$

Untuk estimasi model M-kuantil, diawali dengan menghitung koefisien regresi per kecamatan. Gunakan interpolasi untuk mendapatkan q_{jd} sebagai koefisien unit M-kuantil untuk pasangan (y_{jd}, X_{jd}) dari sampel yang memenuhi persamaan $\hat{y}_{jd} = X_{jd}^T \hat{\beta}_\psi(q_{jd})$ di mana $y_{jd}, X_{jd} \in S_d$. Kemudian, hitung koefisien M-kuantil area sesuai rumus:

$$\hat{\theta}_d = n_d^{-1} \sum_{j \in s_d} q_{jd} \quad (8)$$

Hitung estimasi $\hat{\beta}_\psi$ untuk koefisien M-kuantil kecamatan ($\hat{\theta}_d$) pada persamaan $\hat{y}_{jd} = \mathbf{X}_{jd}^T \hat{\beta}_\psi(\hat{\theta}_d)$ melalui *Iterative Resampling Least Square* (IRLS) dengan mengikuti tahapan berikut:

1. Dimulai iterasi ke-0, hitung estimasi awal $\hat{\beta}^{(0)}$ (koefisien regresi ke-0) menggunakan MKT.
2. Dengan memanfaatkan $\hat{\beta}^{(0)}$, dihasilkan residual $\mathbf{e}_{jd}^{(0)}$ dan nilai skala awal $s^{(0)}$. Penelitian ini menggunakan *Median Absolute Deviation* (MAD) untuk menghitung nilai skala (s).
3. Pilih fungsi pembobot $w(u)$ dan gunakan terhadap scaled residual $\mathbf{u} = \frac{\mathbf{e}_{jd}^{(0)}}{s^{(0)}}$ untuk mendapatkan bobot $\mathbf{w}_{jd}^{(0)}$ awal. Rumus umum w_{jd} adalah $w_{jd} = (1 - \hat{\theta}_d)w(u)$ untuk $\mathbf{e}_{jd} \leq 0$ dan $w_{jd} = \hat{\theta}_d w(u)$ untuk $\mathbf{e}_{jd} > 0$. Bobot awal $\mathbf{w}_{jd}^{(0)}$ menghasilkan matriks diagonal $\mathbf{W}^{(0)}$.
4. Lakukan iterasi ke-1 menggunakan $\mathbf{W}^{(0)}$ untuk mendapatkan estimasi baru $\hat{\beta}^{(1)}$ melalui rumus $\hat{\beta}^{(1)} = (\mathbf{X}^T \mathbf{W}^{(0)} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^{(0)} \mathbf{y}$. Selanjutnya, dari estimasi $\hat{\beta}^{(1)}$, kita dapatkan residual baru $\mathbf{e}_{jd}^{(1)}$. Dari residual baru, kita dapatkan $s^{(1)}$ dan dengan fungsi pembobot, didapatkan bobot $w(u)$. Sampai pada tahap ini, didapatkan matriks diagonal $\mathbf{W}^{(1)}$.
5. Lakukan iterasi ke-2. Dengan menggunakan matriks $\mathbf{W}^{(1)}$ didapatkan estimasi $\hat{\beta}^{(2)}$, $\mathbf{e}_{jd}^{(2)}$, dan $s^{(2)}$ yang baru. Tahap ini menghasilkan matriks diagonal $\mathbf{W}^{(2)}$.
6. Iterasi IRLS menggunakan rumus umum:

$$\hat{\beta}^{(iter)} = (\mathbf{X}^T \mathbf{W}^{(iter-1)} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{W}^{(iter-1)} \mathbf{y} \quad (9)$$

Iterasi akan berhenti saat status konvergen tercapai. Konvergen adalah keadaan ketika perubahan residual lebih kecil dari nilai ε dan didapatkan estimasi $\hat{\beta}_\psi$ akhir sebagai koefisien regresi M-kuantil. Konvergen dinyatakan dengan rumus:

$$\sqrt{\frac{\sum_{jd=1}^{n_d} (e_{jd}^{(iter-1)} - e_{jd}^{(iter)})^2}{\sum_{jd=1}^{n_d} e_{jd}^{(iter-1)^2}} < \varepsilon \quad (10)$$

Di mana ε adalah nilai kecil positif sebesar 0,0001.

Setelah memperoleh $\hat{\beta}_\psi$, hitung residual $\mathbf{e}_{jd} = \mathbf{y}_{jd} - \hat{y}_{jd}$ dan buat model $\hat{y}_{kd} = \mathbf{X}_{kd}^T \hat{\beta}_\psi(\hat{\theta}_d) + \mathbf{e}_{jd}^*$ dimana \mathbf{e}_{jd}^* adalah sampel acak dengan pengembalian berukuran N_d pada residual \mathbf{e}_{jd} , sedangkan \mathbf{X}_{kd} adalah matriks variabel penyerta dari SP2010. Gunakan 50 iterasi simulasi monte carlo untuk menghitung estimasi P_0 dan P_2 berdasarkan pada rumus kemiskinan FGT:

$$\hat{P}_{\alpha,d}^L = \frac{1}{N_d} \sum_{i=1}^{N_d} \left(\frac{t - \hat{y}_{kd}}{t} \right)^\alpha I(\hat{y}_{kd} \leq t) \quad (11)$$

Hitung estimasi P_0 dan P_1 akhir dari rata-rata simulasi sebelumnya, yaitu:

$$\hat{P}_{\alpha,d} = \frac{1}{L} \sum_{l=1}^L \hat{P}_{\alpha,d}^{*L} \quad (12)$$

RMSE dihitung dengan mengadopsi rumus yang diusulkan Marchetti. Dimulai dengan membangkitkan bootstrap populasi U^{*b} , di mana tiap U^{*b} dihitung estimasi $\hat{P}_{\alpha,d}^{*b}$. Kemudian tiap U^{*b} diambil 100 kali bootstrap sampel U^{*r} tanpa pengembalian sebanyak jumlah sampel area kecil $n_d^* = n_d$ untuk menghitung estimasi $\hat{P}_{\alpha,d}^{*br}$. Sedangkan $\tilde{P}_{\alpha,d}^{*br}$ adalah rata-rata $\hat{P}_{\alpha,d}^{*br}$ sesuai persamaan berikut:

$$\bar{P}_{\alpha,d}^{*br} = R^{-1} \sum_{r=1}^R \hat{P}_{\alpha,d}^{*br} \quad (13)$$

$$RRMSE(\hat{P}_{\alpha,d}) = \frac{RMSE(\hat{P}_{\alpha,d})}{\hat{P}_{\alpha,d}} \times 100\% \quad (17)$$

Estimasi bias dan varian sesuai rumus:

$$\hat{B}(\hat{P}_{\alpha,d}) = B^{-1}R^{-1} \sum_{b=1}^B \sum_{r=1}^R (\hat{P}_{\alpha,d}^{*br} - \hat{P}_{\alpha,d}^{*b}) \quad (14)$$

$$\hat{V}(\hat{P}_{\alpha,d}) = B^{-1}R^{-1} \sum_{b=1}^B \sum_{r=1}^R (\hat{P}_{\alpha,d}^{*br} - \bar{P}_{\alpha,d}^{*br})^2 \quad (15)$$

Sehingga rumus RMSE menjadi

$$RMSE(\hat{P}_{\alpha,d}) = \sqrt{\hat{B}(\hat{P}_{\alpha,d})^2 + \hat{V}(\hat{P}_{\alpha,d})} \quad (16)$$

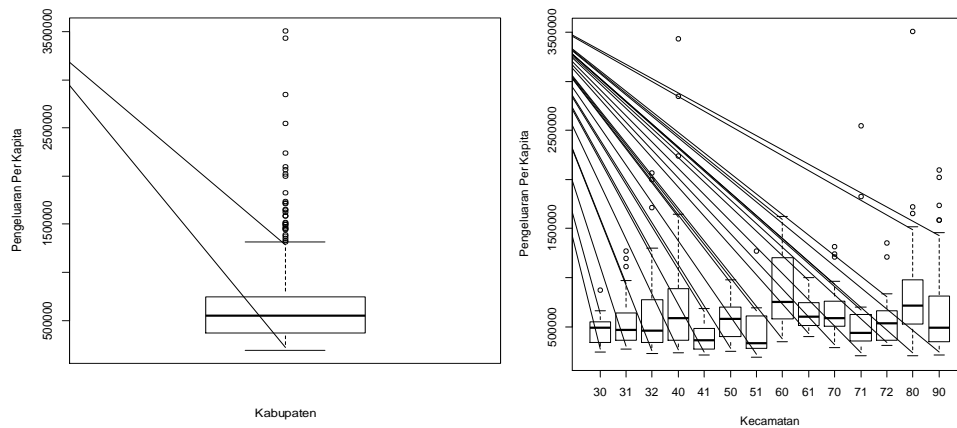
Bandingkan akurasi antara kedua teknik estimasi dengan memakai rumus RRMSE:

HASIL DAN PEMBAHASAN

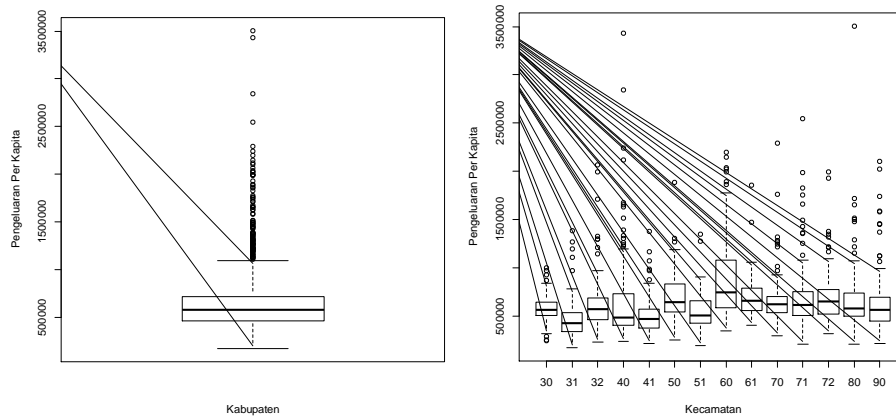
Tabel 4 menunjukkan statistik deskriptif variabel pengeluaran perkapita rumah tangga. Rata-rata dan median dari survei yang sampelnya ditambah (oversampling) berbeda tipis dengan survei yang sampelnya tetap. Simpangan baku pada oversampling lebih rendah dari sampel awal.

Tabel 4.
Pengeluaran Perkapita Rumah Tangga (rupiah)

Data	Mean	Median	Std. Dev	Min	Max
Tanpa Penambahan Sampel	647 751	547 516	407 918	191 097	3 507 652
Dengan Penambahan Sampel	640 368	573 953	313 455	169 313	3 507 652



Gambar 2 *Boxplot* pengeluaran perkapita rumah tangga kabupaten dan kecamatan di Kabupaten Musi Rawas (sampel kecil/awal)



Gambar 3. *Boxplot* pengeluaran perkapita rumah tangga kabupaten dan kecamatan di Kabupaten Musi Rawas (*oversampling*)

Gambar 2 dan 3 memperlihatkan pola sebaran pengeluaran perkapita rumah tangga per bulan dalam satu kabupaten maupun setiap kecamatan. Gambar memperlihatkan bahwa data pengeluaran perkapita rumah tangga per bulan dalam satu kabupaten maupun setiap kecamatan tidak simetris. Pola yang terbentuk menjulur kekanan (*positive skewness*). Hal ini diperkuat oleh adanya pencilan dibagian

atas *boxplot* dan disertai *whisker* bagian atas lebih panjang. Adanya pencilan ini, apabila kita menggunakan metode kuadrat terkecil akan menghasilkan estimasi yang kurang baik. Oleh karena itu penggunaan metode estimasi yang bersifat kekar diperlukan sebagai metode alternatif yang sama baiknya dengan metode kuadrat terkecil tetapi tidak terlalu dipengaruhi oleh pencilan atau hal lain dalam asumsi model.

Tabel 5.
Nilai estimasi koefisien M-kuantil area $\hat{\theta}_d$

Kecamatan	Huber	Hampel
STL Ulu	0.23	0.21
Selangit	0.38	0.36
Sumber Harta	0.42	0.41
Tugumulyo	0.48	0.47
Purwodadi	0.18	0.17
Muara Beliti	0.35	0.33
TP Kepungut	0.38	0.37
Jayaloka	0.70	0.69
Suka Karya	0.56	0.54
Muara Kelingi	0.55	0.53
BTS Ulu	0.32	0.30
Tuah Negeri	0.44	0.42
Muara Lakitan	0.70	0.69
Megang Sakti	0.38	0.37

Tabel 5 menunjukkan nilai estimasi koefisien M-kuantil untuk kecamatan yang dihasilkan dari gugus sampel $(y_{jd}, X_{jd} \in S_d)$. Nilai tersebut merupakan rata-rata koefisien M-kuantil dari unit contoh. Nilai $\hat{\theta}_d$ menggambarkan posisi rata-rata kuantil pengeluaran per kapita berdasarkan data sampel setiap kecamatan. Setelah koefisien M-kuantil kecamatan $(\hat{\theta}_d)$ diperoleh, langkah selanjutnya adalah mencari estimasi β_ψ bersyarat pada nilai koefisien M-kuantil kecamatan $(\hat{\theta}_d)$ menggunakan persamaan $\hat{y}_{jd} = X_{jd}^T \hat{\beta}_\psi(\hat{\theta}_d)$ dimana $y_{jd}, X_{jd} \in S_d$.

$(\hat{\theta}_d)$. Interpretasi hasil tersebut adalah tiap kecamatan memiliki model estimasi regresi linier M-kuantil pada nilai kuantil $\hat{\theta}_d$ tertentu. Semakin tinggi nilai estimasi koefisien M-kuantil area $\hat{\theta}_d$, hasil nilai estimasi β_ψ pada nilai koefisien M-kuantil $\hat{\theta}_d$ makin besar. Nilai estimasi pada variabel penyerta pertama dan kedua yang bernilai negatif menunjukkan keduanya memiliki hubungan berbanding terbalik dengan variabel respon. Nilai estimasi pada variabel penyerta ketiga yang bernilai positif menunjukkan hubungan searah dengan peubah respon.

Tabel 6 dan 7 menampilkan estimasi β_ψ pada nilai koefisien M-kuantil kecamatan

Tabel 6 .
Estimasi β_ψ pada koefisien M-kuantil $\hat{\theta}_d$ dengan fungsi pembobot Huber

Kec	$\hat{\beta}_{0\psi}(\hat{\theta}_d)$	$\hat{\beta}_{1\psi}(\hat{\theta}_d)$	$\hat{\beta}_{2\psi}(\hat{\theta}_d)$	$\hat{\beta}_{3\psi}(\hat{\theta}_d)$
STL Ulu	766992.4	-69194.7	-77671.1	36698.0
Selangit	902718.8	-77876.0	-127736.6	45966.7
Sumber Harta	949235.7	-80600.1	-147522.4	49410.4
Tugumulyo	1009244.0	-84017.6	-172725.6	54251.8
Purwodadi	721620.6	-66436.9	-63296.8	33547.4
Muara Beliti	875280.2	-76131.4	-117157.2	44155.7
TP Kepungut	904642.1	-77994.0	-128524.6	46089.9
Jayaloka	1318612.4	-97977.9	-336633.6	91924.3
Suka Karya	1100052.4	-88402.0	-217727.8	63148.5
Muara Kelingi	1087667.8	-87802.0	-211560.2	61699.0
BTS Ulu	854139.1	-74803.6	-109348.1	42862.3
Tuah Negeri	965230.9	-81518.3	-154313.3	50784.2
Muara Lakitan	1315490.3	-97852.4	-334928.9	91472.8
Megang Sakti	909109.4	-78267.4	-130397.8	46383.1

Tabel 7.

Estimasi β_ψ pada koefisien M-kuantil $\hat{\theta}_d$ dengan fungsi pembobot Hampel

Kec	$\hat{\beta}_{0\psi}(\hat{\theta}_d)$	$\hat{\beta}_{1\psi}(\hat{\theta}_d)$	$\hat{\beta}_{2\psi}(\hat{\theta}_d)$	$\hat{\beta}_{3\psi}(\hat{\theta}_d)$
STL Ulu	754439.9	-68663.4	-63370.5	37139.1
Selangit	901399.0	-77593.8	-120993.4	48083.9
Sumber Harta	952677.5	-80722.1	-142270.8	52142.1
Tugumulyo	1003778.3	-83604.6	-161301.0	56660.7
Purwodadi	711816.0	-65965.0	-49961.5	34139.4
Muara Beliti	867871.7	-75494.3	-107685.8	45507.4
TP Kepungut	915395.3	-78439.8	-126587.0	49126.7
Jayaloka	1309172.6	-99245.3	-312559.4	89596.2
Suka Karya	1074945.8	-87789.2	-191578.8	63456.4
Muara Kelingi	1063944.0	-87157.2	-186649.5	62316.1
BTS Ulu	844751.4	-74076.8	-98428.7	43642.1
Tuah Negeri	961489.4	-81263.5	-145799.5	52924.6
Muara Lakitan	1296307.3	-98722.5	-305126.7	87961.5
Megang Sakti	909048.8	-78047.0	-123893.4	48630.5

Tabel 8 menampilkan hasil estimasi *oversampling* dan estimasi model M-kuantil P_0 menggunakan estimasi langsung dengan P_0 di seluruh kecamatan.

Tabel 8.

Estimasi P_0 tingkat kecamatan (dalam persen)

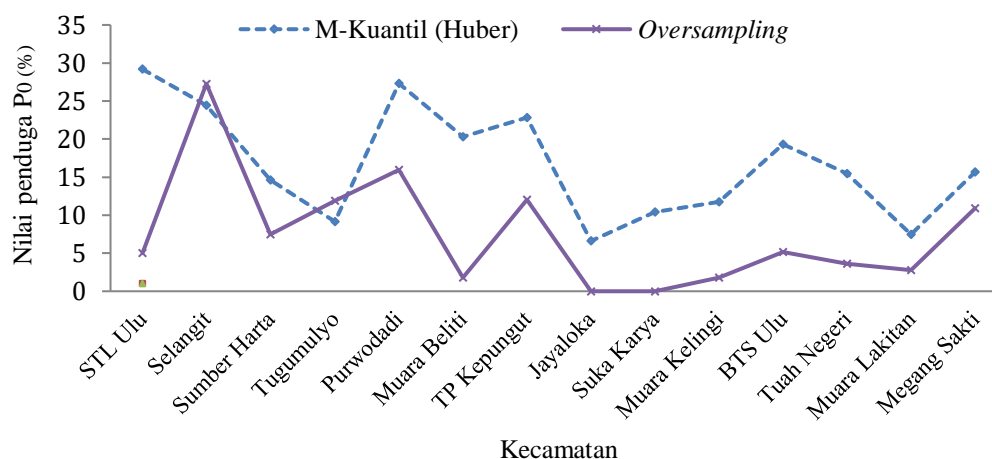
Nama Kecamatan	Estimasi tidak langsung		Estimasi langsung <i>oversampling</i>
	dengan M-kuantil		
	Huber	Hampel	
STL Ulu	29.18	29.93	5
Selangit	24.47	24.08	27.27
Sumber Harta	14.61	14.05	7.5
Tugumulyo	9.16	9.09	11.9
Purwodadi	27.33	27.81	15.97
Muara Beliti	20.28	20.4	1.82
TP Kepungut	22.86	21.67	12
Jayaloka	6.6	6.37	0
Suka Karya	10.43	10.58	0
Muara Kelingi	11.77	12.08	1.82
BTS Ulu	19.29	19.45	5.17
Tuah Negeri	15.48	15.54	3.64
Muara Lakitan	7.47	7.57	2.8
Megang Sakti	15.7	15.37	10.88

Nama Kecamatan	Estimasi tidak langsung dengan M-kuantil		Estimasi langsung <i>oversampling</i>
	Huber	Hampel	
	Rata-rata ^a	18.13	
Median ^a	17.5	17.5	6.34
Level Kab	16.08	15.95	7.63

^a Penghitungan rata-rata dan median tidak melibatkan Kecamatan Jayaloka dan Suka karya

Terlihat pada tabel, bahwa M-kuantil mampu melakukan estimasi angka kemiskinan P_0 yang bernilai lebih dari 0% pada kecamatan Jayaloka dan Suka Karya, yang sebelumnya estimasi langsung P_0 pada dua kecamatan tersebut sebesar 0%. Selain itu, terlihat hasil estimasi P_0 menggunakan

estimasi langsung dengan *oversampling* menghasilkan persentase penduduk miskin lebih rendah dibandingkan M-kuantil. Gambar 6 berikut menampilkan grafik perbandingan estimasi langsung dengan *oversampling* dan M-kuantil fungsi pembobot Huber.



Gambar 4. Perbandingan estimasi P_0 kecamatan

Tabel 9 menunjukkan RRMSE estimasi P_0 pada estimasi langsung *oversampling* dan M-Kuantil. Rata-rata RRMSE estimasi P_0 pada estimasi tidak langsung lebih kecil dari estimasi langsung, artinya M-kuantil

lebih akurat dibandingkan dengan estimasi langsung dengan *oversampling*. Simpangan baku dan median RRMSE P_0 melalui M-kuantil juga lebih kecil dibandingkan dengan estimasi langsung *oversampling*.

Tabel 9.
RRMSE estimasi P_0 tingkat kecamatan (dalam persen)

Nama Kecamatan	Estimasi tidak langsung dengan M-kuantil		Estimasi langsung <i>oversampling</i>
	Huber	Hampel	
STL Ulu	35.23	34.78	39.92
Selangit	18.84	21.6	15.56
Sumber Harta	29.1	26.91	31.88
Tugumulyo	21.07	23.98	20.77
Purwodadi	48.66	48	21.1
Muara Beliti	27.31	22.5	70.7
TP Kepungut	30	32.13	27.29
Jayaloka	73	57.94	-
Suka Karya	50.24	38.38	-
Muara Kelingi	39.26	32.61	70.05
BTS Ulu	26.75	27.35	39.68
Tuah Negeri	24.87	24.33	49.47
Muara Lakitan	43.79	39.65	49.69
Megang Sakti	18.41	20.3	23.38
Rata-rata ^a	30.27	29.51	38.29
Median ^a	28.21	27.13	35.78
Simp baku ^a	9.7	8.28	18.64

^a Penghitungan rata-rata, median, dan simpangan baku tidak melibatkan Kecamatan Jayaloka dan Suka karya

Tabel 10 menunjukkan bahwa M-kuantil mampu mengungkap kemiskinan (P_1) yang tidak lagi bernilai 0% untuk Kecamatan Jayaloka dan Suka Karya.

Tabel 10.
Estimasi P_1 tingkat kecamatan (dalam persen)

Nama Kecamatan	Estimasi tidak langsung dengan M-kuantil		Estimasi langsung <i>oversampling</i>
	Huber	Hampel	
STL Ulu	10.33	10.63	0.68
Selangit	9.37	9.2	3.82
Sumber Harta	4.35	4.15	1.38
Tugumulyo	2.73	2.7	1.77
Purwodadi	8.33	8.52	3.36
Muara Beliti	6.72	6.76	0.39
TP Kepungut	8.35	8.01	1.2
Jayaloka	2.19	2.17	0
Suka Karya	3.18	3.21	0
Muara Kelingi	4.03	4.1	0.19

Nama Kecamatan	Estimasi tidak langsung dengan M-kuantil		Estimasi langsung <i>oversampling</i>
	Huber	Hampel	
BTS Ulu	6.12	6.22	0.73
Tuah Negeri	4.96	4.89	0.16
Muara Lakitan	2.58	2.6	0.7
Megang Sakti	4.65	4.54	1.59
Rata-rata ^a	6.04	6.03	1.33
Median ^a	5.54	5.56	0.97
Level Kab	5.3	5.2	1.16

^a Penghitungan rata-rata dan median tidak melibatkan Kecamatan Jayaloka dan Suka karya

Tabel 11 menunjukkan RRMSE lebih akurat dibandingkan dengan estimasi estimasi P_1 pada estimasi langsung dan M- langsung *oversampling*. Dilihat dari nilai Kuantil. Rata-rata estimasi M-kuantil lebih median, estimasi M-kuantil juga lebih kecil dari estimasi langsung sehingga bisa unggul dibandingkan estimasi langsung. diambil kesimpulan bahwa M-kuantil masih

Tabel 11.
RRMSE estimasi P_1 tingkat kecamatan (dalam persen)

Nama Kecamatan	Estimasi tidak langsung dengan M-kuantil		Estimasi langsung <i>oversampling</i>
	Huber	Hampel	
STL Ulu	40.85	41.32	52.21
Selangit	25.84	29.01	23.27
Sumber Harta	42.73	41.67	38.1
Tugumulyo	34.47	34.48	24.96
Purwodadi	69.5	60.42	22.99
Muara Beliti	36.16	28.56	73.14
TP Kepungut	45.52	43.56	45.19
Jayaloka	72.5	58.88	-
Suka Karya	72.24	56.75	-
Muara Kelingi	48.4	43.44	77.08
BTS Ulu	30.58	33.44	55.32
Tuah Negeri	38.33	38.86	67.56
Muara Lakitan	46.09	41.53	58.17
Megang Sakti	28.81	29.09	28.38
Rata-rata ^a	40.61	38.78	47.2
Median ^a	39.59	40.09	48.7
Simp baku ^a	11.58	8.92	19.74

^a Penghitungan rata-rata, median, dan simpangan baku tidak melibatkan Kecamatan Jayaloka dan Suka karya

Perbandingan Estimasi Langsung dengan Estimasi Tidak Langsung

Secara umum, model M-kuantil menghasilkan rata-rata RRMSE P_0 dan P_1 lebih kecil dibandingkan dengan estimasi langsung *oversampling*. Fungsi pembobot yang menghasilkan RRMSE rata-rata

terkecil untuk P_0 adalah Hampel sebagaimana juga untuk P_1 . Studi empiris telah menunjukkan bahwa M-kuantil layak digunakan sebagai alternatif untuk menggantikan penambahan sampel pada Suseda sekaligus mampu meningkatkan akurasinya.

Tabel 12.
Rata-rata RRMSE estimasi indikator kemiskinan (%)

Teknik Estimasi	HCR	PG	PS
Estimasi langsung			
Dengan Oversampling	38.29	47.2	55.94
Estimasi tidak langsung (M-kuantil)			
Huber	30.27	40.61	48.99
Hampel	29.51	38.78	46.01

SIMPULAN

Pada estimasi langsung, terdapat dua kecamatan yang estimasi indikator kemiskinannya sebesar 0%. Hal ini menunjukkan bahwa penambahan sampel hanya meningkatkan akurasi. Sebaliknya, M-kuantil mampu menghitung estimasi indikator kemiskinan pada beberapa kecamatan yang indikator kemiskinannya 0% jika menggunakan estimasi langsung. Berdasarkan penelitian ini, kami menyimpulkan bahwa M-kuantil mampu

menggantikan survei tambahan dengan penambahan sampel. Penambahan sampel tidak perlu dilakukan lagi karena tidak efisien.

Berdasarkan rata-rata RRMSE, M-kuantil menghasilkan estimasi yang lebih akurat dibandingkan estimasi langsung *oversampling*. Kesimpulan lain dari penelitian ini menunjukkan bahwa fungsi pembobot Hampel lebih baik dari Huber karena memberikan rata-rata RRMSE yang lebih rendah.

DAFTAR PUSTAKA

- Badan Pusat Statistik. (2016). Data dan Informasi Kemiskinan Kabupaten/ Kota (2015). Jakarta: Badan Pusat Statistik.
- Breckling J, Chambers R. (1988). M-quantiles. *Biometrika*. 75(4).761-771.
- Chambers R., Tzavidis N. (2006). M-Quantile Models for small area estimation. *Biometrika*. 93(2).255-268.
- Draper NR, Smith H. (1998). *Applied Regression Analysis Third Edition*. New York (US): Wiley.
- Dunstan R, Chambers R. (1989). Estimating Distribution Functions from Survey Data with Limited Benchmark Information. *Austral. J. Statist.* 31(1).1-11.

- Foster J., Greer J., Thorbecke E. (1984). A class of decomposable poverty measures. *Econometrica*, 52(3).761-766.
- Girinoto. (2017). Kajian pendugaan area kecil untuk indikator kemiskinan melalui pendekatan regresi kekar M-kuantil (studi kasus: tingkat kecamatan pada Kabupaten Bogor) *M.Si thesis*. Institut Pertanian Bogor, Bogor.
- Giusti C., Marchetti S., Pratesi M., Salvati N. (2012). Robust small area estimation and oversampling in the estimation of poverty indicators. *Survey Research Methods*. 6(3).155-163.
- Huber PJ, Ronchetti EM. (2009). *Robust Statistics*. New York (US): John Wiley & Sons.
- Kennickell A.B. (2007). The Role of Oversampling of the Wealthy in the Survey of Consumer Finances. *ISI 56th conference*.
- Koenker R, Bassett G. (1978). Regression quantiles. *Econometrica*. 46(1).33-50.
- Marchetti S., Tzavidis N., Pratesi M. (2012). Non-parametric bootstrap mean squared error estimation for M-quantile estimators of small area averages, quantiles and poverty indicators. *Computational Statistics and Data Analysis*, 56.2889-2902.
- Molina I, Rao J.N.K. (2010). Small area estimation of poverty indicators. *The Canadian Journal of Statistics*. 38(3).369-385.
- Montgomery, DC, Peck, EA, Vining GG. (2012). *Introduction to Linear Regression Analysis Fifth Edition*. New York (US): John Wiley & Sons.
- Rao J.N.K. (2003). *Small Area Estimation*. New York, US: John Wiley and Sons.
- Tzavidis N., Marchetti S., Chambers R. (2010). Robust Estimation of Small Area Means and Quantiles. *Australian and New Zealand Journal of Statistics*. 52(2).167-186.