

ADOPSI REGRESI BEDA UNTUK MENGATASI BIAS VARIABEL TEROMISI DALAM REGRESI DERET WAKTU: MODEL KEHILANGAN AIR DISTRIBUSI DI PDAM SUKABUMI

Yusep Suparman
Universitas Padjadjaran
yusep.suparman@unpad.ac.id

ABSTRAK. Kehilangan air distribusi merupakan salah satu permasalahan yang dihadapi oleh perusahaan air minum daerah, termasuk salah satunya PDAM Sukabumi. Untuk dapat menangani permasalahan ini secara efektif, kita perlu mengetahui besar kontribusi dari setiap faktor yang mempengaruhi kehilangan air ini. Namun demikian, dari sejumlah faktor yang secara konseptual teridentifikasi berkontribusi, tidak semua datanya tersedia. Hal ini dapat mengakibatkan taksiran kontribusi faktor-faktor kehilangan air yang diperoleh dari sebuah model regresi menjadi bias. Dalam makalah ini, saya mengadopsi salah satu pendekatan data panel dalam mengatasi bias variabel teromisi, yaitu regresi beda, ke dalam pemodelan kehilangan air distribusi yang merupakan regresi deret waktu. Dari kasus PDAM Sukabumi, saya menemukan bahwa, regresi beda dapat mengoreksi bias variabel teromisi. Akan tetapi, taksiran galat baku dari taksiran koefisien regresi menjadi relatif lebih besar, sehingga mempengaruhi hasil inferensi koefisien regresinya.

Kata Kunci: *regresi deret waktu; regresi beda; bias variabel teromisi; kehilangan air distribusi.*

1. PENDAHULUAN

Semakin menurunnya kualitas lingkungan akibat dari semakin berkurangnya hutan dan meningkatnya pencemaran berdampak pada kualitas dan kuantitas sumber air bersih alami. Sumur, sungai dan danau yang selama ini menjadi sumber air alami, banyak yang menjadi mudah kering dan kualitas airnya menurun sehingga tidak layak untuk dipergunakan. Dalam kondisi seperti ini, peranan air ledeng menjadi sangat penting untuk pemenuhan kebutuhan keluarga. Selain merupakan mode penyediaan air bersih termurah, air ledeng merupakan mode yang paling efektif untuk menangani penyakit-penyakit yang terkait dengan tingkat sanitasi dan higien yang rendah [5].

Pada 2015 hanya 16,7% dari keluarga Indonesia yang dilayani oleh sambungan air ledeng [2]. Kondisi ini tidak jauh dari tahun-tahun sebelumnya. Pertumbuhan persentase keluarga yang dilayani oleh layanan air ledeng sangat lambat. Dalam lima tahun terakhir hanya naik sebesar 7,1%, [1]. Kondisi ini tidak terlepas dari kemampuan perusahaan air minum daerah (kabupaten/kota) untuk mengembangkan jaringan layanannya. Kemampuannya terkendala, salah satunya oleh kerugian perusahaan yang diakibatkan oleh kehilangan air pada saat pendistribusian. Sebagai contohnya Perusahaan Daerah Air Minum (PDAM) Kota Sukabumi. Setiap tahunnya antara dua sampai empat juta meter kubik air hilang dalam pendistribusian. PDAM Sukabumi pun telah melakukan upaya-upaya penanggulangan.

Dari sekian upaya yang sudah dilakukan, belum ada kajian yang memberikan evaluasi tentang efektivitasnya. Untuk itu perlu diketahui kontribusi dari upaya-upaya tersebut. Hal ini dapat kita lakukan melalui sebuah model regresi kehilangan air yang di dalamnya, kita dapat menaksir kontribusi dari faktor penyebab dan upaya penanganan.

Namun demikian, terkait kehilangan air, tidak semua faktornya dapat kita ukur. Misalkan, banyaknya kebocoran di pipa distribusi utama sangat sulit untuk menghitungnya. Contoh lainnya, tidak adanya data tentang banyaknya sambungan liar. Jika kita mengabaikan faktor-faktor ini dan ternyata mereka berkorelasi dengan faktor lainnya atau upaya yang ada dalam model regresi, dampaknya kualitas taksiran kontribusi dari faktor dan upaya di dalam model menjadi rendah; taksiran koefisien regresi menjadi bias [3].

Bias akibat dari variabel teromisi biasanya ditangani dengan menggunakan pendekatan variabel instrumental [3]. Untuk kasus kehilangan air, sangat sulit untuk menemukan variabel instrumental yang tepat. Selain itu, asumsi dalam pendekatan variabel instrumental tidak dapat kita uji, sehingga kita tidak dapat melakukan validasi hasil. Dalam penelitian ini saya mengusulkan untuk mengadopsi regresi beda [5] sebagai alternative penanganan bias variabel teromisi pada regresi deret waktu. Regresi beda merupakan pendekatan untuk mengatasi bias variabel teromisi dalam model regresi panel yang tentunya berlain dengan model regresi kehilangan air yang merupakan model regresi deret waktu.

2. MODEL KEHILANGAN AIR DISTRIBUSI

Data yang dipergunakan dalam penelitian ini adalah data bulanan yang dikumpulkan oleh PDAM Sukabumi dari Januari 2004 sampai Juni 2009. Variabel terikat dalam model ini adalah volume kehilangan air bulanan dalam meter kubik (y_t). Sedangkan variabel penjelasnya adalah:

- Banyak kebocoran yang dilaporkan selama satu bulan dalam unit (x_{1t}),
- Banyak meter air tidak siap baca dalam unit (x_{2t}),
- Banyak meter air tidak siap baca diganti pada bulan sebelumnya dalam unit (x_{3t-1}),
- Banyak pipa distribusi direhab dalam unit (x_{4t}),
- Banyak penutupan sambungan liar (x_{5t-1}).

Dua variabel penjelas pertama merupakan faktor yang berkontribusi positif terhadap kehilangan air distribusi. Tiga variabel berikutnya merupakan variabel yang terkait dengan upaya PDAM dalam menangani kehilangan air, sehingga kontribusinya bersifat negatif. Perhatikan bahwa selain kelima variabel ini masih terdapat variabel lain seperti, banyaknya kebocoran yang tidak terdeteksi dan banyaknya sambungan liar yang seharusnya masuk ke dalam model. Hal inilah yang kita antisipasi untuk menghindari bias taksiran dari parameter regresi terkait kelima variabel tersebut. Secara formal model regresi kehilangan air distribusi dapat dirumuskan sebagai

$$y_t = \beta_0 + \beta_1 x_{1t} + \beta_2 x_{2t} + \beta_3 x_{3t-1} + \beta_4 x_{4t} + \beta_5 x_{5t-1} + \varepsilon_t \quad (1)$$

Perhatikan, ketika banyaknya kebocoran dan banyaknya sambungan liar tidak masuk kedalam model, keduanya akan teragregasi ke dalam kekeliruan model ε_t . Dan ketika keduanya berkorelasi dengan setidaknya salah satu dari x , kekeliruan model ε_t akan berkorelasi dengan x . Dalam hal ini asumsi $E(\varepsilon|\mathbf{x})=0$ tidak terpenuhi dan taksiran ordinary least square untuk model tersebut menjadi bias [3]. Untuk itu, kita rumuskan regresi beda untuk persamaan (1) sebagai berikut:

$$(y_t - y_{t-1}) = \beta_1^*(x_{1t} - x_{1t-1}) + \beta_2^*(x_{2t} - x_{2t-1}) + \beta_3^*(x_{3t-1} - x_{3t-2}) + \beta_4^*(x_{4t} - x_{4t-1}) + \beta_5^*(x_{5t-1} - x_{5t-2}) + \varepsilon_t^* \quad (2)$$

Dalam kasus model panel pendekatan ini dipergunakan untuk variabel teromisi yang *time-invariant*. Untuk kasus ini secara analisis dapat dibuktikan bahwa taksiran yang diperoleh bersifat tak bias [6]. Untuk kasus kehilangan air, banyaknya kebocoran dan banyaknya sambungan liar merupakan variabel teromisi yang *time-varying*. Namun demikian perubahannya dari waktu ke waktu relatif kecil, sehingga pendekatan regresi beda ini setidaknya akan mengurangi bias variabel teromisi dengan baik.

Asumsi dari model di atas sama halnya dengan asumsi model regresi standar. Sehingga untuk menaksir parameternya saya mengguakan pendekatan *ordinary least square* (OLS) (lihat [6]).

3. HASIL DAN PEMBAHASAN

Untuk menunjukkan adanya bias, saya pergunakan OLS untuk menaksir parameter berdasar model (1) dan model (2). Hasilnya saya sajikan di dalam Tabel 1. Taksiran parameter dari model (1) pada umumnya jauh lebih besar dari pada taksiran parameter dari model (2). Dari kedua model, taksiran dari model (2) lebih realistis dibandingkan dengan taksiran dari model (1). Hal yang paling mencolok adalah taksiran untuk β_3 yaitu kontribusi dari usaha penggantian meter air. Kontribusi dari penggantian meter air bersifat negatif, adanya penggantian meter air dapat mengurangi kehilangan air distribusi, bukannya menambah kehilangan air. Taksiran model (2) sesuai dengan hal ini. Sementara taksiran model (1), bertentangan dengan hal ini.

Taksiran untuk β_1 pun menunjukkan hal yang sama. Taksiran dari model (1) jauh lebih besar dari taksiran dari model (2). Kebocoran yang dilaporkan adalah kebocoran pada pipa distribusi tersier yaitu pada pipa distribusi berukuran kecil yang terhubung dengan pipa pelanggan. Satu kebocoran pada pipa jenis ini tidak akan mengakibatkan kehilangan air sampai tiga ribu meter kubik seperti yang di berikan oleh model (1). Pelaporan kebocoran pada pipa tersier biasanya segera dilaporkan pelanggan karena akan mengganggu pasokan air ke pelanggan tersebut. Selain itu, perbaikan dapat dilakukan dengan cepat. Taksiran dari model (2) yang jauh lebih kecil, lebih dapat diterima dari pada taksiran dari model (1).

Hal serupa juga ditunjukkan oleh taksiran β_5 . Penggunaan rata-rata hampir sampai empat ribu meter kubik oleh satu pelanggan, baik itu instansi maupun industri apalagi rumah tangga, sangat tidak mungkin. Tapi kalau penggunaan dalam satu bulan rata-rata sekitar dua ratus lima puluh meter kubik oleh satu pelanggan, ini masih memungkinkan. Sehingga untuk satu sambungan liar, penutupannya akan mengurangi kehilangan air seperti yang ditaksirkan oleh model (2).

Dari uraian ini, dapat kita ketahui bahwa taksiran model (2) lebih bisa diterima dibandingkan dengan dari model (1). Dengan demikian model (2) dapat memperbaiki bias yang terjadi dalam model (1). Namun demikian, taksiran galat baku dari model (2) relatif lebih besar dibandingkan dari model (1). Kalau kita perhatikan kekeliruan model (2) merupakan fungsi dari kekeliruan dua model regresi, hal ini mengakibatkan, galat baku dari taksirannya relati lebih besar. Dengan demikian, pengujian berdasarkan taksiran dari model (2) akan tidak sensitif terhadap penolakan hipotesis nol. Sehingga diperlukan modifikasi untuk taksiran galat baku model (2) supaya sebanding dengan taksiran galat baku dari model (1). Untuk selanjutnya, saya hanya akan membahas taksiran tanpa melihat signifikansinya.

Tabel 1. Taksiran Parameter

Parameter	Taksiran	
	Model (1)	Model (2)
β_0	78852,9 (35473,6)	t.t. t.t.
β_1	2976,6 (507,0)	491,8 (444,0)
β_2	308,2 (57,9)	165,4 (65,6)
β_3	153,6 (248,8)	-140,6 (255,0)
β_4	-4185,7 (2316,6)	-692,2 (1704,1)
β_5	-3948,7 (2488,4)	-259,4 (1427,6)
R^2	0,77	0,39

Berdasarkan nilai taksiran dari model (2) dapat kita ketahui bahwa dari kedua faktor kebocoran pipa merupakan kontributor terbesar terhadap kehilangan air distribusi. Setiap terjadi kebocoran, rata-rata air distribusi hilang sebanyak 491,8 m³. Sementara itu, kehilangan dari meter air tidak siap baca lebih kecil kontribusinya. Setiap unit meter air tidak siap baca rata-rata akan menambah kehilangan air sebanyak 165,4 m³. Sedangkan upaya pengurangan kehilangan yang paling besar kontribusinya adalah rehab pipa distribusi. Setiap unit pipa yang direhab akan mengurangi kehilangan air sebanyak 692,2m³. Yang berikutnya adalah penutupan sambungan liar. Setiap penutupan sambungan liar rata-rata akan mengurangi kehilangan air 259m³. Dan yang terakhir penggantian satu unit meter air akan mengurangi rata-rata kehilangan 140,6m³.

4. SIMPULAN

Dari taksiran parameter kita dapat mengetahui bahwa taksiran kontribusi dari masing-masing faktor dan upaya. Berdasarkan informasi ini dapat lebih jauh dilakukan analisis untuk menentukan upaya mana yang menurut biaya lebih efektif untuk dilakukan. Sementara itu, dari aspek metodologi, dapat kita ketahui bahwa penggunaan regresi beda dapat mengurangi omitted variable bias pada regresi time series. Dalam hal ini, variabel teromisinya diasumsikan mendekati *time-invariant*. Sedangkan yang perlu diperbaiki dalam pendekatan ini adalah koreksi terhadap taksiran galat bakunya. Selain, regresi beda, ada *constrained autoregression* [4] yang dapat dipergunakan sebagai alternatif. Pendekatan ini dapat dipergunakan baik untuk variabel teromosi yang *time-varying* maupun *time-invariant*. Namun demikian, diperlukan metode kemungkinan maksimum untuk penaksiran parameternya. Selain itu ada pembatasan nilai-nilai parameter yang tidak dapat dilakukan oleh program-program pengolahan data standar. Dari aspek praktis regresi beda lebih mudah untuk dikerjakan.

DAFTAR PUSTAKA

- [1] BPS. 2011. *Statistik Kesejahteraan Rakyat 2010*. Jakarta: Badan Pusat Statistik.
- [2] BPS. 2016. *Statistik Kesejahteraan Rakyat 2015*. Jakarta: Badan Pusat Statistik.
- [3] Greene, W.H. (2008). *Econometrics analysis* (5th ed.). New Jersey: Prentice Hall.
- [4] Suparman, Y., Folmer, H., & Oud, J. H. L. (2014). Hedonic price models with omitted variables and measurement errors: a constrained autoregression-structural equation modeling approach with application to urban Indonesia. *Journal of Geographical Systems*, 16(1), 49-70. DOI: 10.1007/s10109-013-0186-3
- [5] WHO-UNICEF (2013) WHO/UNICEF joint monitoring program for water supply and sanitation: Data and estimates. WHO <http://www.wssinfo.org/data-estimates/table/> (04/21/2013)
- [6] Wooldridge, J.M. (2002). *Introductory econometrics: A modern approach* (2nd ed.). South-Western.